

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification 6 : C12N 15/12, C12Q 1/00, A61K 45/00, 38/17, C07K 14/705, C12N 15/70, 1/21 // (C12N 1/21, C12R 1:19)</p>		<p>A2</p>	<p>(11) International Publication Number: WO 97/07209</p> <p>(43) International Publication Date: 27 February 1997 (27.02.97)</p>
<p>(21) International Application Number: PCT/GB96/01980</p> <p>(22) International Filing Date: 14 August 1996 (14.08.96)</p> <p>(30) Priority Data: 9517098.1 21 August 1995 (21.08.95) GB</p> <p>(71) Applicant (<i>for all designated States except US</i>): IMPERIAL COLLEGE OF SCIENCE TECHNOLOGY AND MEDICINE [GB/GB]; Sherfield Building, Exhibition Road, London SW7 2AZ (GB).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (<i>for US only</i>): LITTLE, Peter [GB/GB]; Imperial College of Science Technology and Medicine, Biochemistry Dept., Prince Consort Road, London SW7 2BY (GB). HADJANTONAKIS, Anna-Katerina [GB/GB]; Imperial College of Science Technology and Medicine, Biochemistry Dept., Prince Consort Road, London SW7 2BY (GB).</p> <p>(74) Agents: HARDING, Charles, Thomas et al.; D. Young & Co., 21 New Fetter Lane, London EC4A 1DA (GB).</p>		<p>(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p>Published</p> <p><i>Without international search report and to be republished upon receipt of that report.</i></p> <p><i>With an indication in relation to a deposited microorganism furnished under Rule 13bis separately from the description.</i></p> <p><i>Date of receipt by the International Bureau:</i> 9 September 1996 (09.09.96)</p>	
<p>(54) Title: RECEPTOR</p> <p>(57) Abstract</p> <p>A method of screening a chemical for subsequent use as a pharmaceutical agent. The method comprises contacting the chemical with a receptor, and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises EGF-like repeats and/or cadherin-like repeats.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

RECEPTOR

The present invention relates to a receptor. In particular the present invention relates to the use of a receptor to screen agents to assess their suitability for subsequent use 5 as pharmaceutical agents, such as therapeutic agents and diagnostic agents.

Receptors are structures that bind chemical stimuli specifically and directly or indirectly transduce a message into the intracellular environment (Watson *et al* 1992 Recombinant DNA Second Edition, Chapter 17, published by Scientific American 10 Books). Some receptors, otherwise known as G-protein coupled receptors (GCRs), are coupled to second messenger systems *via* GTP-binding proteins, otherwise known as G-proteins. The G-proteins connect hormone receptors to adenylate cyclase or other signalling enzymes.

15 In more detail, the GCRs represent the largest receptor protein class in vertebrates. Typically they are seven-pass transmembrane receptors (see Figure 1). In particular, the GCRs have been shown to be involved in the regulation of a large variety of physiological processes, with many of the genes encoding them mutated in genetic disorders and mutants.

20 GCRs, can be divided into six families on the basis of their amino acid sequence homologies (similarities) which span across the transmembrane containing region. In this regard, the B family, which is the second largest family, contains the receptors for pituitary adenylate cyclase activating polypeptide, vasoactive intestinal polypeptide 25 (VIP), secretin, growth hormone releasing hormone, diuretic hormone, glucagon, glucagon-like peptide, calcitonin and gastric inhibitory polypeptide.

GCRs can even be placed into functional categories. In this regard, several different types of G-protein coupling have been identified.

For example, the GCR will either interact with an ion channel which is itself a seven-pass transmembrane protein or it can interact with adenylate cyclase, phospholipase C or phospholipase A2, all of which signal to secondary messengers.

5 The G-protein can either be stimulatory (Gs) or inhibitory (Gi) and can therefore stimulate or inhibit the action of the ion channel or second messenger pathway they are effecting.

10 All the B family GCRs have been shown to couple to adenylate cyclase *via* a stimulatory G-protein.

In this regard, the present invention provides a new receptor obtainable from animals. The present invention also provides a new use of that receptor.

15 Thus, according to a first aspect of the present invention there is provided a method of screening a chemical for subsequent use as a pharmaceutical agent, the method comprising contacting the chemical with a receptor; and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises EGF-like repeats.

20 According to a second aspect of the present invention there is provided a method of screening a chemical for subsequent use as a pharmaceutical agent, the method comprising contacting the chemical with a receptor; and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises cadherin-like repeats.

25 According to a third aspect of the present invention there is provided a method of screening a chemical for subsequent use as a pharmaceutical agent, the method comprising contacting the chemical with a receptor; and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises EGF-like repeats and cadherin-like repeats.

Preferably, the method includes contacting the chemical-receptor complex with a G-protein and determining whether the chemical-receptor complex stimulates the G-protein.

5 Preferably the receptor resembles or is a GCR.

In the following commentary, the term "receptor according to the present invention" includes the receptor as defined in the above-mentioned aspects of the present invention.

10

According to a fourth aspect of the present invention there is provided a method of affecting neural development comprising stimulating a receptor with a chemical; wherein the receptor is the receptor according to the present invention. This method can be an *in vitro* or an *in vivo* method.

15

According to a fifth aspect of the present invention there is provided the use of the receptor according to the present invention to screen chemicals for subsequent use as a pharmaceutical.

20

According to a sixth aspect of the present invention there is provided a chemical that has been screened by the method of the present invention.

25

According to a seventh aspect of the present invention there is provided a receptor capable of interacting with a G-protein and wherein the receptor comprises EGF-like repeats.

According to a eighth aspect of the present invention there is provided a protein comprising the amino acid sequence represented as SEQ. I.D. No. 1, or a fragment, homologue or variant thereof.

30

According to a ninth aspect of the present invention there is provided a protein comprising the amino acid sequence represented as SEQ. I.D. No. 3, or a fragment, homologue or variant thereof.

5 According to a tenth aspect of the present invention there is provided a nucleotide sequence comprising the nucleotide sequence represented as SEQ. I.D. No. 2, or a fragment, homologue or variant thereof, and/or wherein the nucleotide sequence is obtainable from one or more of NCIMB No. 40766, NCIMB No. 40767 and NCIMB No. 40768.

10

According to an eleventh aspect of the present invention there is provided a nucleotide sequence comprising the nucleotide sequence represented as SEQ. I.D. No. 4, or a fragment, homologue or variant thereof, and/or wherein the nucleotide sequence is obtainable from one or more of NCIMB No. 40766, NCIMB No. 40767 and NCIMB No. 40768.

15

According to an twelfth aspect of the present invention there is provided a vector capable of expressing the receptor according to the present invention or the protein according to the present invention, or comprising the nucleotide sequence according 20 to the present invention.

According to a thirteenth aspect of the present invention there is provided a construct comprising or capable of expressing any one of the vector according to the present invention, the receptor according to the present invention, the protein according to the 25 present invention, or the nucleotide sequence according to the present invention.

According to a fourteenth aspect of the present invention there is provided a cell, tissue or organ comprising or capable of expressing any one of the construct according to the present invention, the vector according to the present invention, the receptor 30 according to the present invention, the protein according to the present invention, or the nucleotide sequence according to the present invention.

According to a fifteenth aspect of the present invention there is provided an organism comprising or capable of expressing any one of the cell, tissue or organ according to the present invention, the construct according to the present invention, the vector according to the present invention, the receptor according to the present invention, the 5 protein according to the present invention, or the nucleotide sequence according to the present invention.

According to a sixteenth aspect of the present invention there is provided a receptor capable of interacting with a G-protein and wherein the receptor comprises EGF-like 10 repeats, wherein the receptor is obtainable from neural tissue and/or ectodermal cells.

According to a seventeenth aspect of the present invention there is provided a receptor capable of interacting with a G-protein and wherein the receptor comprises cadherin-like repeats, wherein the receptor is obtainable from neural tissue and/or ectodermal 15 cells.

According to an eighteenth aspect of the present invention there is provided a receptor capable of interacting with a G-protein and wherein the receptor comprises EGF-like repeats and cadherin-like repeats, wherein the receptor is obtainable from neural tissue 20 and/or ectodermal cells.

According to a nineteenth aspect of the present invention there is provided an assay kit comprising a surface having attached thereto or contained within or on any one of the organism according to the present invention, the cell, tissue or organ according to 25 the present invention, the construct according to the present invention, the vector according to the present invention, the receptor according to the present invention, the protein according to the present invention, or the nucleotide sequence according to the present invention.

30 Typically the assay kit will comprise a series of titre wells capable of holding a suitable sample of the present invention (e.g. the receptor or the gene coding for same

in cells or in a cell free environment). Preferably, the assay kit comprises a series of titre wells, wherein at least one of which well holds a suitable sample of the present invention (e.g. the receptor or the gene coding for same in cells or in a cell free environment). Optionally, the assay kit may comprise one or more G-proteins.

5

As mentioned above, if the assay kit of the present invention comprises the receptor of the present invention then that assay kit would be useful for screening chemicals that are capable of interacting with the receptor to form a chemical-receptor complex. With that assay kit, the interaction of the chemical-receptor complex with the G-10 protein can be observed either directly or indirectly. An example of indirect observation is observing a change (e.g. an increase) in cAMP levels.

Alternatively, if the assay kit of the present invention comprises the nucleotide sequence of the present invention then that assay kit would be useful for screening 15 chemicals for affecting expression of that sequence.

Other aspects of the present invention include the use of the receptor of the present invention to screen for agents that are capable of any one or more of
20 stimulating the receptor to cause neural cells to divide;
stimulating the receptor to cause neural cells to differentiate;
stimulating the receptor to affect cellular physiology; and
25 stimulating the receptor for use in the repair of trauma and/or in the treatment of neurodegenerative diseases.

Other aspects of the present invention include the use of the receptor of the present invention for one or more of:

- 5 stimulating adenylate cyclase;
- increasing cAMP levels; and
- promoting neural growth.

These uses can be *in vitro* or *in vivo* uses.

10 Other aspects of the present invention include NCIMB No. 40766, NCIMB No. 40767, and NCIMB No. 40768.

Further aspects of the present invention include:

15 a pharmaceutical preparation consisting of or comprising the receptor of the present invention, optionally admixed with or contained within a suitable carrier, diluent or excipient.

20 a pharmaceutical preparation consisting of or comprising the nucleotide sequence of the present invention, optionally admixed with or contained within a suitable carrier, diluent or excipient.

25 a pharmaceutical preparation consisting of or comprising a chemical when screened by the method of the present invention, optionally admixed with or contained within a suitable carrier, diluent or excipient.

the use of the receptor according to present invention in the manufacture of a medicament to treat neural disorder.

30 the use of the nucleotide sequence according to present invention in the manufacture of a medicament to treat neural disorder.

the use of a chemical when screened by the method of the present invention in the manufacture of a medicament to treat neural disorder.

5 A further aspect of the present invention includes a method of treating a subject in need of, or likely to be in need of, treatment for neural disorder wherein the method comprises administering to the subject a receptor according to the present invention, or a protein expressed by the nucleotide sequence according to the present invention, or a chemical when screened by the method of the present invention.

10 An additional aspect of the present invention includes a hybrid receptor, and genes coding for the same and vectors etc. comprising same, wherein the hybrid receptor comprises at least a part of the receptor of the present invention and at least a part of another receptor, such as another receptor or even part or all of a G-protein. The hybrid receptor is advantageous as it allows one to affect and/or to tailor the 15 stimulation of the receptor to one or more stimuli.

Preferably the receptor is obtainable from neural tissue.

Preferably the receptor is obtainable from ectodermal cells.

20 Preferably the receptor is prepared by use of recombinant DNA techniques.

Preferably the receptor is obtainable from deposit NCIMB No. 40766 and/or deposit NCIMB No. 40767.

25 Alternatively, the receptor is obtainable from deposit NCIMB No. 40768.

Preferably the receptor comprises the amino acid sequence represented as SEQ. I.D. No. 1, or is a fragment, homologue or variant thereof.

Preferably the receptor comprises the amino acid sequence represented as SEQ. I.D. No. 3, or is a fragment, homologue or variant thereof.

5 Preferably the receptor is expressed by the nucleotide sequence represented as SEQ. I.D. No. 2, or is a fragment, homologue or variant thereof.

Preferably the receptor is expressed by the nucleotide sequence represented as SEQ. I.D. No. 4, or is a fragment, homologue or variant thereof.

10 Preferably the chemical is screened to determine if it is useful for one or more of:

- i. causing neural cells to divide;
- ii. causing neural cells to differentiate;
- iii. affecting cellular physiology;
- 15 iv. repairing trauma;
- v. treating neurodegenerative diseases;
- vi. stimulating adenylate cyclase production;
- vii. increasing cAMP levels;
- viii. promoting neural growth.

20

Preferred embodiments of the present invention therefore include:

- i. a receptor capable of interacting with a G-protein and comprising EGF-like repeats and/or cadherin-like repeats, wherein the receptor is obtainable from 25 neural tissue and/or ectodermal cells;
- ii. a receptor comprising the sequence represented as SEQ. I.D. No. 1, or a fragment, homologue or variant thereof capable of being stimulated by a chemical;

30

- iii. a receptor comprising the sequence represented as SEQ. I.D. No. 3, or a fragment, homologue or variant thereof capable of being stimulated by a chemical;
- 5 iv. a receptor encoded by the nucleotide sequence represented as SEQ. I.D. No. 2, or a fragment, homologue or variant thereof and wherein the expressed protein is capable of being stimulated by a chemical;
- 10 v. a receptor encoded by the nucleotide sequence represented as SEQ. I.D. No. 4, or a fragment, homologue or variant thereof and wherein the expressed protein is capable of being stimulated by a chemical;
- 15 vi. a receptor obtainable from deposit NCIMB No. 40766;
- vii. a receptor obtainable from deposit NCIMB No. 40767; and
- 20 viii. a receptor obtainable from deposit NCIMB No. 40768.

With this aspect of the present invention, the receptor may comprise a plurality of any combination of the features i. to viii. as listed above.

In a highly preferred embodiment the receptor of the present invention is not expressed by the natural genomic DNA sequence when in its natural environment. Thus, highly preferred embodiments include the receptor when prepared by use of recombinant 25 DNA techniques, including the expression of cDNA or a synthetic nucleotide sequence.

Preferably the receptor is expressed by a cDNA sequence that is obtainable from any one or more of deposits NCIMB 40766, NCIMB 40767 and NCIMB 40768.

In addition, or alternatively, preferably the receptor is expressed by a cDNA sequence that comprises sequence shown as SEQ. I.D. No. 2, or a sequence that is complementary thereto, or a sequence having substantial homology therewith, or a sequence containing any suitable codon substitutions but wherein the resultant protein
5 is capable of acting as receptor as herein defined.

In addition, or alternatively, more preferably the receptor is expressed by a cDNA sequence that comprises sequence shown as SEQ. I.D. No. 4, or a sequence that is complementary thereto, or a sequence having substantial homology therewith, or a
10 sequence containing any suitable codon substitutions but wherein the resultant protein is capable of acting as receptor as herein defined.

In a highly preferred embodiment the nucleotide sequence coding for the receptor of the present invention is not in its natural environment and under the control of the
15 promoter with which it is naturally associated which is also in its natural environment.

Thus, highly preferred embodiments include the use of recombinant DNA techniques using for example cDNA or a synthetic nucleotide sequence.

20 In a highly preferred embodiment the nucleotide sequence coding for the receptor of the present invention is not in its natural environment and under the control of the promoter with which it is naturally associated which is also in its natural environment, wherein the receptor comprises the amino acid sequence shown as SEQ.I.D. No. 1, more preferably SEQ.I.D. No. 3, or variant, fragment or homologue thereof, wherein
25 the nucleotide sequence is a cDNA sequence that is obtainable from any one or more of deposits NCIMB 40766, NCIMB 40767 and NCIMB 40768, and wherein the nucleotide sequence comprises the sequence shown as SEQ. I.D. No. 2, more preferably SEQ. I.D. No. 4, or a sequence that is complementary thereto, or a sequence having substantial homology therewith, or a sequence containing any suitable
30 codon substitutions but wherein the nucleotide sequence codes for a protein that is capable of behaving as a receptor as herein defined.

Other embodiments of the present invention include: a transformed host organism (such as a microorganism, such as *E. coli*) capable of producing the receptor according to the present invention as a consequence of the introduction of a nucleotide sequence as herein described; a method for preparing the receptor according to the 5 present invention comprising expressing a nucleotide sequence according to the present invention contained in the host organism and isolating the expressed receptor; and a vector (such as a transformed *pBLUESCRIPT* plasmid, *pGEX* plasmid or *pCDNA3* plasmid) incorporating the nucleotide sequence according to the present invention. By way of example, the receptor can be expressed in *E. coli*, baculovirus, yeast or 10 mammalian cell expression systems.

All of the above-mentioned aspects of the present invention optionally include the combination of the receptor of the present invention with a G-protein.

15 The term "chemical" includes any chemical compound, including nucleotide sequences both in sense and antisense orientation, proteins, enzymes etc. The term also includes a ligand, wherein a ligand is a natural substance that naturally binds to the receptor.

20 The term "pharmaceutical agent" includes chemicals for use as diagnostic and/or therapeutic purposes. The term also includes pharmaceutical agents for human and/or veterinary applications.

25 The terms "screen" and "screening" include the use of the receptor according to the present invention to screen agents to assess their suitability for subsequent use as pharmaceutical agents, such as therapeutic agents and diagnostic agents.

30 The term "receptor" is used in its normal sense as typically meaning a protein that spans the membrane of a cell and that can bind, on its extra-cellular side, a chemical (otherwise known as a ligand). Binding of the chemical causes changes to the receptor that result in a chemical (enzymatic) reaction being initiated on the intra-cellular part of the receptor. These changes are the first part of a signalling chain of actions that

result in some change to the cells physiology. In the case of GCRs, binding of the chemical causes the receptor to effect a G-protein with which it is associated on the inner membrane surface. This disturbance results in changes to the enzymatic state of the G-protein, which then interacts with the signalling system.

5

The term "G-protein" is used in its normal sense as typically meaning a protein which is associated with a receptor and which is capable of being effected by the receptor. Changes in the receptor (binding of chemicals/ligands) effect the enzymatic state of the G-protein and these changes can effect the interactions of the G-protein with other 10 proteins which are components of a cascade of signalling events. The outcome of the signalling is dependent on the nature of the cell containing receptor and G-protein. In a preferred embodiment, the receptor interacts with a G-protein.

15 The term "EGF-like repeats" is used in its normal sense as typically meaning a protein sequence similar to the following "consensus" sequence.

$CX_{2-6}CX_{4-6}CX_{5-10}CXCX_{8-22}C$

wherein C is cysteine and X is any amino acid.

20 Preferably, the receptor of the present invention comprises at least one EGF-like repeat and/or at least one cadherin-like repeat. Typically, the receptor of the present invention has between 1 and/or 36 EGF-like repeats and between 1 to 36 cadherin-like repeats. Preferably, the receptor of the present invention has between 1 and 20 EGF-like repeats and/or between 1 and 20 cadherin-like repeats. Preferably, the receptor 25 of the present invention has between 3 and 10 EGF-like repeats and/or between 3 and 10 cadherin-like repeats. More preferably, the receptor of the present invention has between 4 and 7 EGF-like repeats and/or between 6 and 10 cadherin-like repeats. More preferably, the receptor of the present invention has at least about 6 EGF-like repeats and/or at least about 10 cadherin-like repeats.

30

Preferably, the receptor of the present invention comprises at least one EGF-like repeat and at least one cadherin-like repeat.

Preferably, the receptor of the present invention has between 1 and 36 EGF-like repeats and between 1 to 36 cadherin-like repeats. Preferably, the receptor of the present invention has between 1 and 20 EGF-like repeats and between 1 and 20 cadherin-like repeats. Preferably, the receptor of the present invention has between 3 and 10 EGF-like repeats and between 3 and 10 cadherin-like repeats. More preferably, the receptor of the present invention has between 4 and 7 EGF-like repeats and between 6 and 10 cadherin-like repeats. More preferably, the receptor of the present invention has at least about 6 EGF-like repeats and at least about 10 cadherin-like repeats.

The term "chemical-receptor complex" includes binding of the chemical to the receptor, such as by hydrogen bonding and/or covalent bonding. The chemical-receptor complex may then interact with a G-protein. Determination of the formation of the chemical-receptor complex can be achieved by conventional techniques. However, it is preferred to determine formation of the chemical-receptor complex by observing the effect of the complex on a G-protein, such as by observing an increase 20 in cAMP levels.

The term "obtainable from" includes directly or indirectly obtaining the receptor. Examples of indirectly obtaining the receptor include expressing the receptor cDNA by means of a suitable expression system.

25 The terms "variant", "homologue" or "fragment" include any substitution of, variation of, modification of, replacement of, deletion of or addition of one or more amino acid(s)/nucleic acid from or to the sequence providing the resultant protein is capable of behaving as a receptor as herein defined.

The expression "substantial homology", which can be otherwise expressed as "substantial similarity", includes homology with respect to structure and/or nucleotide components, providing the resultant protein is a receptor as herein defined.

- 5 With respect to sequence homology (i.e. similarity), preferably there is at least 50 % homology, preferably at least 60% homology, more preferably at least 75% homology, more preferably at least 80% homology, more preferably at least 85% homology, more preferably at least 90% homology, such as at least 95% homology.
- 10 The term "complementary" means that the present invention also covers recombinant nucleotide sequences that can hybridise to the recombinant nucleotide sequences.

The term "construct" - which is synonymous with terms such as "conjugate", "cassette" and "hybrid" - includes all or part of the nucleotide sequence according to the present invention directly or indirectly attached to another nucleotide sequence, such as a promoter.

The construct may even contain or express a marker which allows for the selection of the genetic construct in the host into which it has been transferred.

20

The construct of the present invention preferably comprises a promoter.

The term "promoter" is used in the normal sense of the art, e.g. an RNA polymerase binding site in the Jacob-Monod theory of gene expression.

25

The term "vector" includes an expression vector and a transformation vector. The term "expression vector" means a construct capable of *in vivo* or *in vitro* expression. The term "transformation vector" means a construct capable of being transferred from one species to another - such as from an *E.Coli* plasmid to another host.

30

The terms "cell", "tissue" and "organ" include cell, tissue and organ *per se* and when within an organism.

5 The term "organism" in relation to the present invention includes any organism that could comprise the recombinant nucleotide sequence coding for the receptor according to the present invention and/or products obtained therefrom, and/or wherein the recombinant nucleotide sequence according to the present invention can be expressed when present in the organism.

10 Preferably the organism is a transgenic organism. The term "transgenic organism" in relation to the present invention includes any organism that comprises the recombinant nucleotide sequence coding for the receptor according to the present invention and/or products obtained therefrom, and/or wherein the recombinant nucleotide sequence according to the present invention can be expressed within the organism. Preferably 15 the recombinant nucleotide sequence is incorporated in the genome of the organism.

The term "protein" includes un-modified and modified proteins such as post-translationally modified proteins and glycosylated proteins.

20 The receptor of the present invention is sometimes referred to as the ME2 protein.

The following samples were deposited in accordance with the Budapest Treaty at the recognised depositary The National Collections of Industrial and Marine Bacteria Limited (NCIMB) at 23 St Machar Drive, Aberdeen, Scotland, AB2 1RY, United 25 Kingdom, on 18 August 1995:

E. coli Xl-1 blue containing mouse cDNA plasmid ME2(22) which was allocated deposit number NCIMB 40766;

E. coli Xl-1 blue containing mouse cDNA plasmid ME2(78) which was allocated deposit number NCIMB 40767;

30 *E. coli* DH1 containing human cosmid ME2HC20 which was allocated deposit number NCIMB 40768.

These deposits are discussed later in the Experimental Section (see Deposits).

Thus, highly preferred embodiments of the present invention include any one of the aforementioned aspects of the present invention but wherein the receptor or the 5 nucleotide sequence coding for same is obtainable from any one or more of deposits NCIMB 40766, NCIMB 40767 and NCIMB 40768.

The present invention will now be described only by way of examples, in which reference shall be made to the following Figures, in which:

10

Figure 1 is a pictorial representation of a typical GCR;

Figure 2 is a DNA map of the receptor of the present invention;

15

Figure 3 is a pictorial representation of the receptor of the present invention;

Figure 4 is a schematic representation of expression patterns;

Figure 5 is a restriction map;

20

Figure 6 presents SEQ ID No. 1;

Figure 7 presents SEQ ID No. 2;

25

Figure 8 presents SEQ ID No. 3;

Figure 9 presents SEQ ID No. 4; and

30

Figure 10 presents an analysis of an amino acid sequence (sequence range 1 to 2707).

In some of the following commentary, sections (parts) of the gene coding for the receptor of the present invention are sometimes referred to as ME2(x) wherein "x" represents a particular numbered fraction.

5 ME2 Genetic Mapping

Our initial studies revealed that the receptor of the present invention is coded by a single copy gene. This single copy gene is conserved in organisms as diverged as human, mice and fruit flies. In particular, the single copy gene maps to human 10 chromosome region 22^{qter} and mouse chromosome 15. In both of these genomes the receptor gene is contained in a region associated with gastrulation and neural mutants and disorders.

Expression Behaviour

15

To determine the *in vivo* expression of the receptor of the present invention, both reverse transcriptase polymerase chain reaction (RT-PCR) and wholmount *in situ* hybridisation were carried out on wild-type mouse embryos.

20 RT-PCR analysis showed that the receptor of the present invention is first expressed in the early postimplantation embryo between 4 and 6 days post coitum (dpc), then continues until adulthood.

25 The embryonic expression of the receptor of the present invention precedes the start of gastrulation, the event which results in the generation of the three germ layers of the developing embryo. Prior to this, the embryo does not contain neural tissue. Embryonic expression of the gene coding for the receptor of the present invention correlates with cells of ectodermal origin, which go on to form the nervous system.

30

For ease of reference, Figure 4 is a schematic representation of the expression patterns of the receptor of the present invention in the developing central nervous system, in particular in the developing spinal cord (Figure 4(A)) and the developing hindbrain (Figure 4(B)). In this regard, interesting features of this dynamic expression pattern 5 include the delineation of segments in the developing hindbrain and neural tube. In the hindbrain novel sub-rhombomeric expression was observed. In the neural tube, initially the transcripts are ubiquitous and then resolve into 5,4 and finally 2 dorsoventrally restricted domains. one in the roof plate and one in the floor plate. Gene expression is highly localised and persists throughout development. with adult 10 transcripts localised to the brain and eye.

Expression Discussion

Observing the pattern of expression of the receptor of the present invention indicates 15 that it may be involved in the control of neural development. In this regard, the receptor is expressed almost exclusively in neural tissue (which is discussed in more detail later). In particular, expression precedes the first obvious neural structures in the developing embryo and the pattern of subsequent expression is complex. In later embryos, expression around the ventricle of the brain is significant since this is 20 believed to be the region that contains the neural stem cells.

Further observations revealed detection of the receptor of the present invention peri-ventricularly in adult brains. This particular pattern of synthesis is therefore under complex spatial and temporal controls and is the period in which the nervous system 25 is proliferating most rapidly.

Hence, the expression evidence strongly suggests that the receptor of the present invention might play an important part in the "control machinery" of neural development (see later discussion).

Isolation Of The Receptor Of The Present Invention

In some of the following commentary, sections (parts) of the gene coding for the receptor of the present invention are sometimes referred to as ME2(x) wherein "x" 5 represents a particular numbered fraction. The isolation of the complete the receptor of the present invention coding sequence is shown in Figure 5.

In more detail, in step 1 (Figure 5) a mouse 8.5 dpc whole embryo cDNA was screened using a human cDNA clone (16FB2). 16FB2 was originally isolated from a 10 human fetal brain cDNA library by hybridisation with human cosmid ZnFP16 as described in Hoovers *et al.*, (Genomics 10 254-263).

ME2(2) was then isolated from the initial screening of the mouse cDNA library (Figure 5). Extensive sequence analysis of both ME2(2) and 16FB2 has shown that 15 they have sequence homology in a G-rich region in the 3' untranslated region.

Furthermore, complete nucleotide sequencing of ME2(2) showed that it had no homology to any sequences in the publicly accessible DNA databases.

20 ME2(2) was then used immediately in whole mount *in situ* expression analyses and produced the striking expression pattern of the receptor of the present invention. The remainder of the gene sequence coding for the receptor of the present invention was then isolated as follows.

25 In step 2 (Figure 5) ME2(2) was used to re-screen the mouse 8.5dpc cDNA library leading to the isolation of 6 different clones, the largest of which being ME2(22). Sequence analysis and database searches with the 3.2kb ME2(22) sequence identified a large open reading frame whose predicted amino acid sequence had homology to the family B group of G-protein coupled receptors. In this regard, ME2(22) covers the 30 region from the polyA tail to trans-membrane region IV.

In step 3 (Figure 5) the 5' EcoRI to *Pst*I fragment of ME2(19) was used to rescreen which led to the isolation of 3 further cDNA clones.

5 In step 4 (Figure 5) a primer PLKH20 corresponding to sequence 3527 to 3541 in the sequence shown as SEQ.I.D. No. 2 and derived from ME2(42) was used to isolate the H1 fragment using the RACE method (Frohman, M.A. (1993) Methods in Enzymol. 218 340-56). The DNA sequence of H1 was used to identify the DNA from 3295 to 3312 (again see SEQ. I.D. No. 2) which was used to make a new primer, PLKH26, for RACE analysis which gave rise R12 in step 5 (Figure 5).

10

In step 6 (Figure 5), R12 was used to screen the cDNA library which gave rise to, amongst other clones, the cDNA clone ME2(78) which extends almost to the 5' end.

15

The minimal set of cDNA clones that defines all of the receptor of the present invention is ME2(78) and ME2(22) (see Figure 2 and Figure 5).

DNA Analysis

20

The DNA sequence of the receptor of the present invention was established entirely using published methods.

25

In particular, the sequencing methodology used was the Sanger technique (Sanger *et al.*, 1977 Proc. Natl. Acad. Sci. USA. 74 5463-7). The sequencing kits used were supplied by Pharmacia Biotech and the manufacturer's protocols were followed throughout.

30

DNA was sequenced either by analysis of cloned molecules using sequencing primers specific for vector sequences and sequencing into the cDNA, or by synthesising specific primers, obtained from conventional commercial synthesis companies, and using these to establish DNA sequence directly from internal parts of the cloned cDNA molecule.

The sequence data obtained is shown in the attached Figures as SEQ. I.D. No. 2 (see Figure 7) and SEQ. I.D. No. 4 (see Figure 9). A map of the DNA sequence is represented in Figure 2.

5 Re-isolation Of The Receptor Of The Present Invention

The receptor of the present invention was re-isolated by using PCR techniques.

10 Since the gene coding for the receptor of the present invention is over 4.5 kb it is preferable to isolate a cDNA containing the receptor coding regions by a multi-step process, rather than by a one step process using RT-PCR to isolate the whole cDNA. Hence, by using the complete coding sequence for the receptor of the present invention it is possible to isolate a series of cDNA fragments that can then be ligated.

15 In this regard, primers flanking pairs of unique restriction enzyme sites were used to amplify individual regions and subsequent restriction enzyme digestion and ligation to generate the complete sequence. Details of this approach are given below and in Figure 2.

20 Fragment 1: Primer pair(3') PLKH31+(5') PLKH52

Amplification products are digested with *Bsp*H_I.

Fragment 2: Primer pair(3') PLKH47+(5')PLKH59c

Amplification products are digested with *Bsp*H_I

25

Primer	Sequence (5'-3')	Position
PLKH59C	5' CAG CGG GGA CTA CTG CGA GAC TGA AAT	1-27
PLKH47	5' AGC TTG TCG AAG ATG TCA AC	2675-2694
PLKH52	5' ATC TTA CAG CAT GAG AGC CGC C	2414-2435
30 PLKH31	5' GGT AAT GAC ACA GTC ACT GGC ATG	4856-4879

Fragments 1 and 2 were independently amplified from mouse embryonic or adult brain reverse transcribed cDNA under standard PCR conditions. The amplification products were then subsequently directly restriction endonuclease digested with BspHI to give ragged ends. Fragments 1 and 2 were then ligated to each other and then cloned into 5 a T vector.

Construction Of A Complete cDNA Clone

Construction of a complete cDNA clone for the gene for the receptor of the present 10 invention was as follows.

In particular, construction of a complete cDNA clone for the gene for the receptor of the present invention was achieved using the protocol detailed in the previous section.

15 In more detail, the method used relied on the minimal set of cDNA clones obtained from the mouse 8.5dpc libraries mentioned above (see also Figure 2 and Figure 5).

The minimal set of clones ME2(22) and ME2(78) have *EcoRI* linkers and are cloned 20 into pBluescript plasmid vector. Since there is no *EcoRI* site in the ME2 transcript, construction of a complete cDNA clone will be done by *EcoRI* + *AvrII* digestion of ME2(22), isolation of the 2.6kb fragment and ligation of this to the 4.3kb *EcoRI* plus *AvrII* fragment from ME2(78): this product is cloned into the appropriate *EcoRI* digested vector.

25 Deposits

As mentioned above three deposits have been made in accordance with the Budapest Treaty. In this regard, NCIMB 40766 is an *E. coli* Xl-1 blue host with a *pBluescript* 30 SK+ vector containing fragment ME2(22) - i.e. nucleotides 3657 to 6794 (see SEQ.I.D.No. 2).

NCIMB 40767 is an *E. coli* Xl-1 blue host with a *pBluescript* SK+ vector containing fragment ME2(78) - i.e. nucleotides 1 to 4813 (see SEQ.I.D. No. 2).

NCIMB 40768 is a recombinant cosmid containing the main part of the human receptor gene according to the present invention. In this regard, the cosmid vector is *pCos2EMBL* and the host cell *E. coli* DH1. The human DNA derives from a region of human chromosome 22^{qter} and contains parts of the human receptor of the present invention gene including parts of the 7TM region but not extending further 5' than this.

10

In order to prepare a full length cDNA clone from NCIMB 40766 and NCIMB 40767, the appropriate cDNA fractions can be excised by use of suitable restriction enzymes, isolated and then ligated. The full length cDNA can then be inserted into any suitable expression system and subsequently expressed by suitable means. The *pBluescript* SK+ plasmids can be recovered from bacterial cells grown in L-broth containing 100 (μg/ml Ampicillin using routine methods detailed in Sambrook *et al.*, (1989 Molecular cloning - A laboratory manual. Second edition. Cold Spring harbor laboratory press.). Then the receptor cDNA fragments may be isolated subsequent to their excision with *EcoRI*.

20

Likewise, the DNA from NCIMB 40768 and/or fragments thereof can be excised by use of suitable restriction enzymes, isolated, inserted into any suitable expression system and subsequently expressed by suitable means. The DNA can be recovered by growing the bacteria in L-broth supplemented with 30 (μg/ml Kanamycin and recovering the DNA according to routine methods detailed in Sambrook *et al.*, (1989 Molecular cloning - A laboratory manual. Second edition. Cold Spring harbor laboratory press). The human DNA fragments can be resolved by cleavage with almost any 6-base recognition enzyme.

30

Amino acid analysis

The amino acid sequence data are listed in the attached sequence listings as SEQ. I.D. No. 1 (see Figure 6) and SEQ. I.D. No. 3 (see Figure 8).

5

Amino acid analysis of the receptor of the present invention reveals that it appears to be a large membrane spanning receptor having an unusual structure. This structure is pictorially shown in Figure 3.

10 In slightly more detail, the C-terminal region appears to have the structure of a 7-pass transmembrane receptor related to the B family of G-protein coupled receptors (GCRs). Thus it is believed that the receptor of the present invention is a new protein.

15 Further amino acid analysis reveals that the receptor of the present invention contains EGF-like repeats. In this regard, towards the N-terminus (extracellular) of the receptor there are a number of EGF-like repeats (see Figure 3). One of these EGF-like repeats is ~300 amino acids away from the C-terminal sequence. The EGF-like repeats are shown in Figure 10 (marked "EGF 1" etc.). Divergent EGF-like repeats are also marked.

20

A large number of molecules containing EGF-like repeats have been identified in both vertebrates and invertebrates. These molecules include, for example, blood clotting factors and proteins that are required for correct embryonic development.

25 Examples of proteins that are required for correct embryonic development, which molecules have primarily been characterised in invertebrates, include fibropellin, a cell coat protein of sea urchins, glp-1 and lag-12 proteins required for inductive interactions in the nematode worm *C. elegans*, and a number of *Drosophila* proteins including Crumbs, which is required for establishing epithelial cell polarity.

30 Additional examples include Notch, Delta and Serrate proteins, which are required for neurogenesis, and Slit, which is a protein involved in axonal pathfinding. Notch and

its ligands Delta and Serrate are involved in cell-cell signalling that determines adjacent cell fates: though this signal is not directly mitogenic.

5 To date, there have been reports of some isolated vertebrate proteins that have some homology (similarity) to the invertebrate proteins. For example, three Notch genes, a single Delta, Jagged and Delta-like have been identified to date.

10 Other vertebrate proteins that have been isolated are EMR-1 (Baud et al 1995 Genomics 26 334-344) and CD97 (Genebank accession no. X84700). However, there is no mention of possible utility of such proteins. let alone mention of pattern of expression/synthesis.

15 Further amino acid analysis reveals that the receptor of the present invention contains cadherin-like repeats. These cadherin-like repeats are shown in Figure 10 (marked as "CD 1" etc.). Cadherin-like repeats have been implicated in protein-protein interactions (Geiger and Ayalon 1992 Ann Rev Cell Biol 8 307-332).

20 Some of the transmembrane portions of the receptor of the present invention are shown in Figure 10 (marked as "TM 1" etc.).

25 When the amino acid sequence of the receptor of the present invention is compared with the amino acid sequences of known proteins that are required for correct embryonic development it is observed that there is some sequence homology, though this is less than 80%. More importantly, however, in distinction to the receptor of the present invention those proteins are all single-pass transmembrane proteins with the cluster of EGF-like repeats in their extracellular domains. In contrast, the receptor of the present invention has a seven-pass transmembrane topology, similar to a GCR.

30 Accordingly, as there have been no reports in the literature for a receptor found in neural tissue that is capable of interacting with a G-protein but, in addition, having EGF-like repeats so the receptor of the present invention is novel.

Functions Of The Receptor Of The Present Invention

The EGF-like repeats and the 7 pass transmembrane (7TM) structures of the receptor of the present invention suggest important functions and utilities for the receptor. In 5 this regard, the EGF-like repeats will bind a ligand which may, as in the case of Notch/Delta, be a protein attached to another cell, or it may be free, as in the case of blood clotting factors. In either case, binding of a ligand will cause the cytoplasmic region to signal to the cellular machinery, *via* a G-protein. Since the 7TM region is most similar to the family B GCRs regions, it is likely that it will signal *via* a 10 stimulatory G-protein which stimulates adenylate cyclase and causes cAMP levels to increase as all family B receptors operate in this fashion.

Likewise, the cadherin-like repeats of the receptor of the present invention suggest important functions and utilities for the receptor.

15 Without wishing to be bound by theory, it is believed that the outcome of this signalling, based upon known examples of GCRs, could be due to one or more of the following effects:

20 1. Stimulation of the receptor of the present invention might cause neural cells to divide.

2. Stimulation of the receptor of the present invention might cause neural cells to differentiate.

25 3. Stimulation of the receptor of the present invention might cause changes to cellular physiology.

Uses

Based on the above-mentioned functions of the receptor of the present invention, it is clear that the receptor can be used in a number of useful utilities. Some of those 5 utilities are now presented.

1. The receptor of the present invention could be a therapeutic target. In this regard, if the ligand or a modified ligand can be defined, then artificial treatment of neural tissue with this ligand could trigger the receptor of the present invention to signal. 10 This signal would then trigger the normal response of the receptor of the present invention, which would be a way of modulating the growth, function or properties of brain cells that express the receptor of the present invention. This would therefore have an application in the repair of trauma and in treatment of neurodegenerative diseases, such as Parkinson's disease and Alzheimer's disease.
- 15 2. The receptor of the present invention can be used as a test reagent to identify the ligand. This can be done for example in artificial test systems where the receptor of the present invention is expressed from a recombinant expression vector in cells that otherwise do not express the receptor of the present invention. These cells can then 20 be used to test for binding of the ligand by studying cAMP level changes upon treatment of cells with proteins or other cells.
- 25 3. The nucleotide sequence coding for the receptor of the present invention gene could be used to modify the behaviour of cells by transgenesis. Potential areas of application include the modification of cells used for transplantation treatments of degenerative diseases and the modification of whole animals by normal transgenic procedures. Both of these applications would result in cells with modified growth potential.

Screening Protocol

Two screening protocols are now presented.

5 The first is based on Lutz *et al.* (1993. FEBS Letters 334, 3-8). In outline, a ME2 expression vector (for example pcDNA-1) is constructed and then introduced into Cos-7 cells by transfection. This results in the cells expressing ME2 and, because of the biological properties of G-proteins, a G-protein becomes naturally associated with ME2 on the cell surface. The cells are then treated *in vitro* with proteins, chemicals, 10 other cells (intact or broken up). Then one assays for changes to cAMP levels which are caused by ME2 binding a ligand and signalling via the G-protein to alter cAMP levels. If changes are seen, this implies that the ligand is or is contained in, in the substance that was treated with cells. This is the assay of choice for purification of 15 the ligand.

15 The second method is based upon Cheng & Flanagan (1994, Cell 79 157-168). In this regard, one synthesises in *E. coli* and isolates the N-terminal fragment of ME2 (N terminus of ME2 to the membrane entry point in fig 3) which has been fused at this point to the enzyme alkaline phosphatase. This hybrid protein binds to its normal 20 ligand. Binding is then detected by looking for the alkaline phosphatase dragged along at its end. This could be used to isolate cDNA clones containing the normal ME2 ligand using exactly the methods detailed in Cheng & Flanagan (1994, Cell 79 157-168).

25 The present invention therefore relates to a novel receptor and a novel use of that receptor.

The nature of the receptor of the present invention, its spatiotemporally restricted expression, coupled with the evolutionary conservation of the receptor gene suggests 30 that the receptor of the present invention plays a role in the determination of ectodermal cell types within the developing embryo. This has important consequences

as it enables possible pharmaceutical agents to be screened to see if they stimulate the receptor and if so then those agents could be used to promote neural growth. In addition, the receptor can be inserted (such as by way of transplantation or by way of transgenesis of the coding gene) into a subject either in need of treatment or to 5 develop *in vivo* screening techniques.

Other modifications of the present invention will be apparent to those skilled in the art without departing from the scope of the invention.

CLAIMS

1. A method of screening a chemical for subsequent use as a pharmaceutical agent, the method comprising contacting the chemical with a receptor; and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises EGF-like repeats.
2. A method of screening a chemical for subsequent use as a pharmaceutical agent, the method comprising contacting the chemical with a receptor; and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises cadherin-like repeats.
3. A method of screening a chemical for subsequent use as a pharmaceutical agent, the method comprising contacting the chemical with a receptor; and determining whether the chemical interacts with the receptor to form a chemical-receptor complex; wherein the receptor comprises EGF-like repeats and cadherin-like repeats.
4. A method according to any one of claims 1 to 3 wherein the receptor is obtainable from neural tissue.
5. A method according to any one of claims 1 to 4 wherein the receptor is obtainable from ectodermal cells.
6. A method according to any one of the preceding claims wherein the receptor is prepared by use of recombinant DNA techniques.
7. A method according to any one of the preceding claims wherein the receptor is obtainable from deposit NCIMB No. 40766 and/or deposit NCIMB No. 40767.
- 30 8. A method according to any one of claims 1 to 6 wherein the receptor is obtainable from deposit NCIMB No. 40768.

9. A method according to any one of the preceding claims wherein the receptor comprises the amino acid sequence represented as SEQ. I.D. No. 1, or is a fragment, homologue or variant thereof.

5 10. A method according to any one of the preceding claims wherein the receptor comprises the amino acid sequence represented as SEQ. I.D. No. 3, or is a fragment, homologue or variant thereof.

10 11. A method according to any one of the preceding claims wherein the receptor is expressed by the nucleotide sequence represented as SEQ. I.D. No. 2, or is a fragment, homologue or variant thereof.

15 12. A method according to any one of the preceding claims wherein the receptor is expressed by the nucleotide sequence represented as SEQ. I.D. No. 4, or is a fragment, homologue or variant thereof.

13. A method according to any one of the preceding claims wherein the method includes contacting the chemical-receptor complex with a G-protein and determining whether the chemical-receptor complex stimulates the G-protein.

20 14. A method according to any one of claims 1 to 13 wherein the chemical is screened to determine if it is useful for one or more of:

25 i. causing neural cells to divide;
ii. causing neural cells to differentiate;
iii. affecting cellular physiology;
iv. repairing trauma;
v. treating neurodegenerative diseases;
vi. stimulating adenylate cyclase production;

30 vii. increasing cAMP levels;
viii. promoting neural growth.

15. A method of affecting neural development comprising stimulating a receptor with a chemical; wherein the receptor is the receptor as defined in any one of claims 1 to 14.
- 5 16. Use of a receptor as defined in any one of claims 1 to 14 to screen for agents that are capable of stimulating the receptor to cause neural cells to divide.
- 10 17. Use of a receptor as defined in any one of claims 1 to 14 to screen for agents that are capable of stimulating the receptor to cause neural cells to differentiate.
18. Use of a receptor as defined in any one of claims 1 to 14 to screen for agents that are capable of stimulating the receptor to affect cellular physiology.
19. Use of a receptor as defined in any one of claims 1 to 14 to screen for agents 15 that are capable of stimulating the receptor for use in the repair of trauma and/or in the treatment of neurodegenerative diseases.
20. Use of a receptor as defined in any one of claims 1 to 14 to stimulate adenylate cyclase.
- 20 21. Use of a receptor as defined in any one of claims 1 to 14 to increase cAMP levels.
22. Use of a receptor as defined in any one of claims 1 to 14 to promote neural 25 growth.
23. Use of a receptor as defined in any one of claims 1 to 14 to screen chemicals for subsequent use as a pharmaceutical.
- 30 24. A chemical when screened by the method according to any one of claims 1 to 15.

25. A receptor capable of interacting with a G-protein and wherein the receptor comprises EGF-like repeats, wherein the receptor is obtainable from neural tissue and/or ectodermal cells.
- 5 26. A receptor capable of interacting with a G-protein and wherein the receptor comprises cadherin-like repeats, wherein the receptor is obtainable from neural tissue and/or ectodermal cells.
- 10 27. A receptor capable of interacting with a G-protein and wherein the receptor comprises EGF-like repeats and cadherin-like repeats, wherein the receptor is obtainable from neural tissue and/or ectodermal cells.
28. A receptor according to any one of claims 25 to 27 wherein the receptor is the receptor as defined in any one of claims 4 to 14.
- 15 29. A protein comprising the amino acid sequence represented as SEQ. I.D. No. 1, or a fragment, homologue or variant thereof.
30. A protein comprising the amino acid sequence represented as SEQ. I.D. No. 3, or a fragment, homologue or variant thereof.
- 25 31. A nucleotide sequence comprising the nucleotide sequence represented as SEQ. I.D. No. 2, or a fragment, homologue or variant thereof, and/or wherein the nucleotide sequence is obtainable from one or more of NCIMB No. 40766, NCIMB No. 40767 and NCIMB No. 40768.
- 30 32. A nucleotide sequence comprising the nucleotide sequence represented as SEQ. I.D. No. 4, or a fragment, homologue or variant thereof, and/or wherein the nucleotide sequence is obtainable from one or more of NCIMB No. 40766, NCIMB No. 40767 and NCIMB No. 40768.

33. A vector capable of expressing the receptor as defined in any one of claims 1 to 14 or the protein of claim 29 or 30, or comprising the nucleotide sequence as defined in claim 31 or 32.

5 34. A construct comprising or capable of expressing any one of the vector of claim 33, the receptor as defined in any one of claims 1 to 14 or the protein of claim 29 or 30, or comprising the nucleotide sequence as defined in claim 31 or 32.

10 35. A cell, tissue or organ comprising or capable of expressing any one of the construct of claim 34, the vector of claim 33, the receptor as defined in any one of claims 1 to 14 or the protein of claim 29 or 30, or comprising the nucleotide sequence as defined in claim 31 or 32.

15 36. An organism comprising or capable of expressing any one of the cell, tissue or organ of claim 35, the construct of claim 34, the vector of claim 33, the receptor as defined in any one of claims 1 to 14 or the protein of claim 29 or 30, or comprising the nucleotide sequence as defined in claim 31 or 32.

20 37. An assay kit comprising a surface having attached thereto or contained within or on any one of an organism according to claim 36, the cell, tissue or organ of claim 35, the construct of claim 34, the vector of claim 33, the receptor as defined in any one of claims 1 to 14 or the protein of claim 29 or 30, or comprising the nucleotide sequence as defined in claim 31 or 32.

25 38. NCIMB No. 40766.

39. NCIMB No. 40767.

40. NCIMB No. 40768.

41. A method substantially as described herein.
42. A use substantially as described herein.
- 5 43. A receptor substantially as described herein.
44. An amino acid sequence substantially as described herein.
45. A nucleotide sequence substantially as described herein.

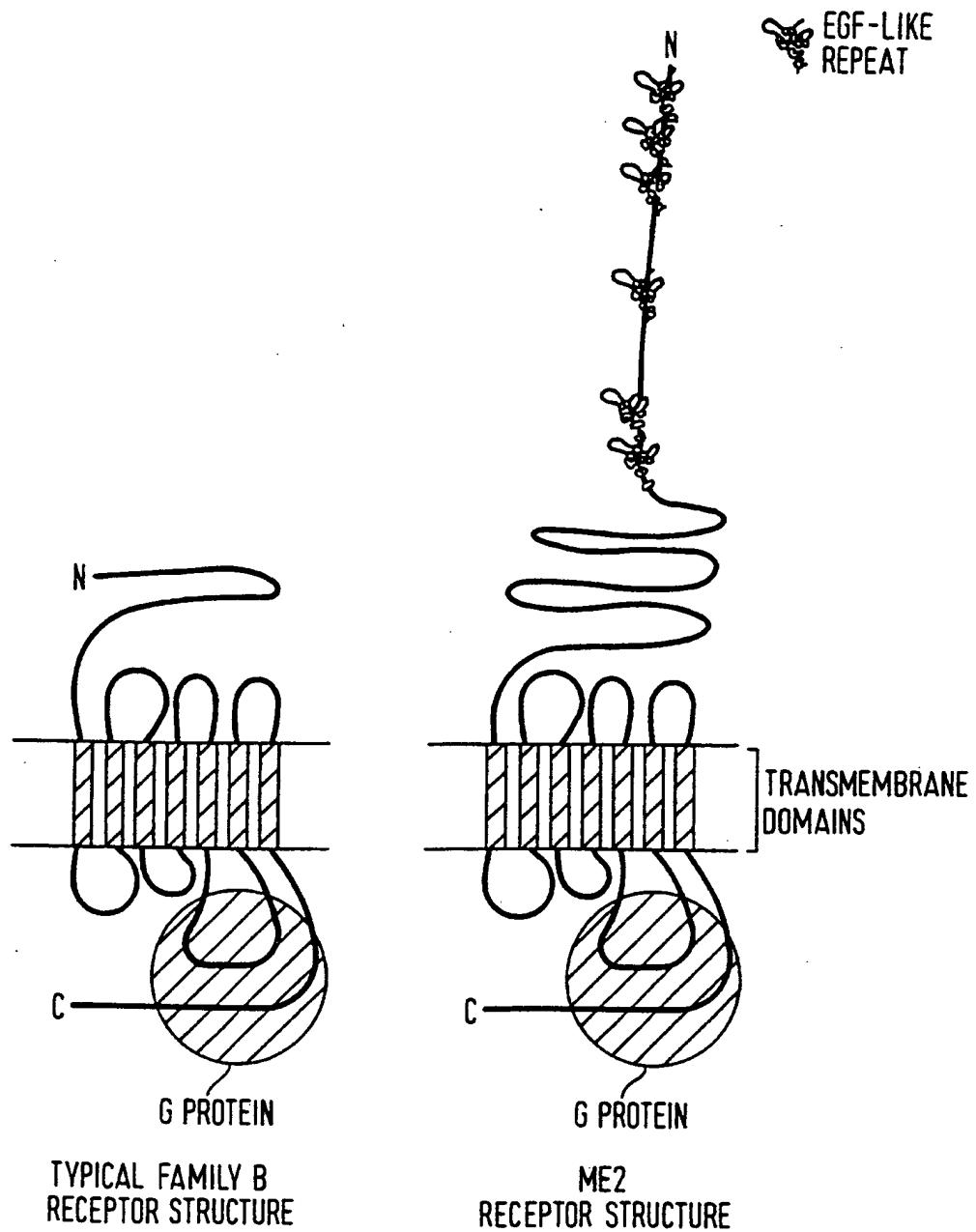
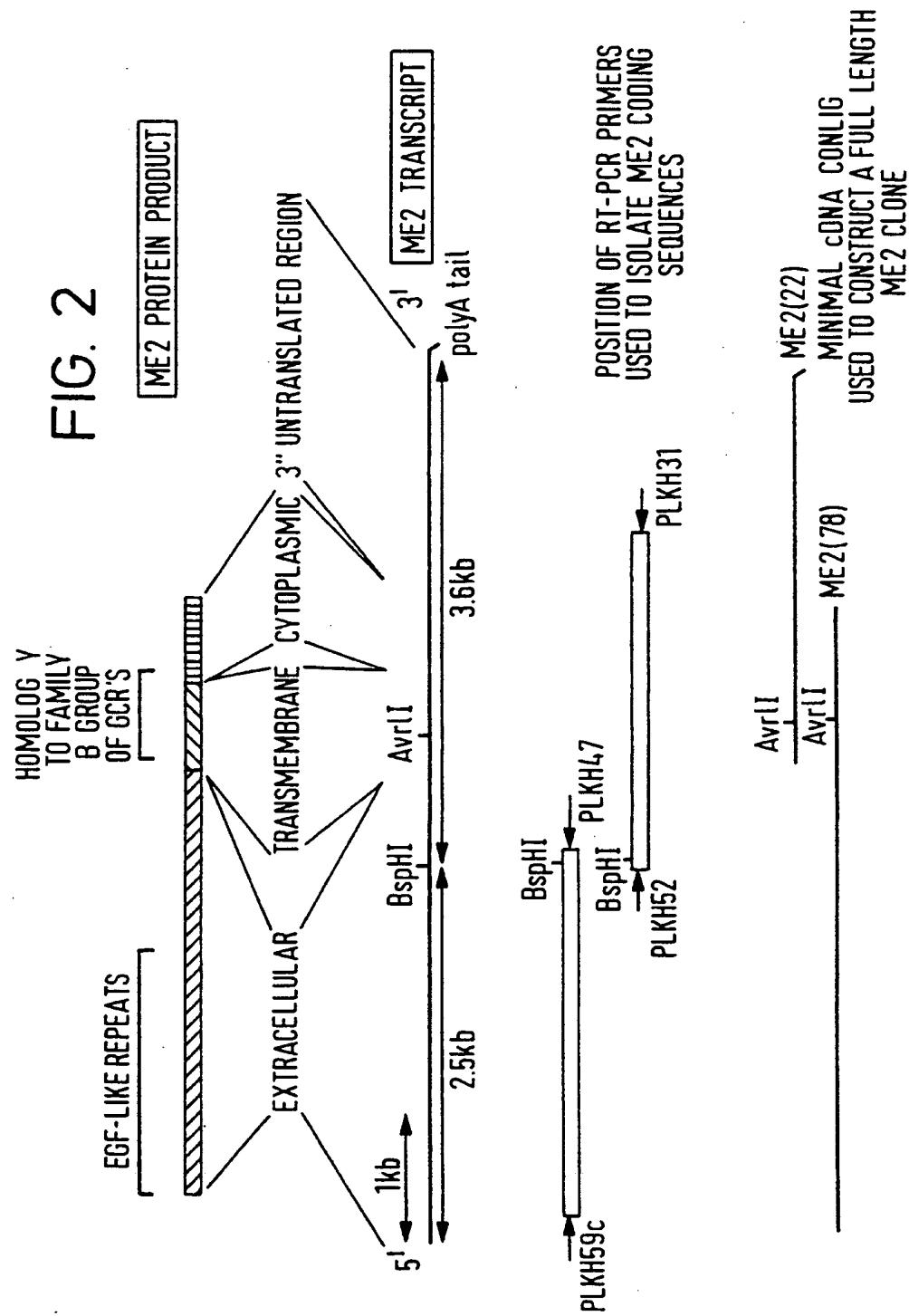


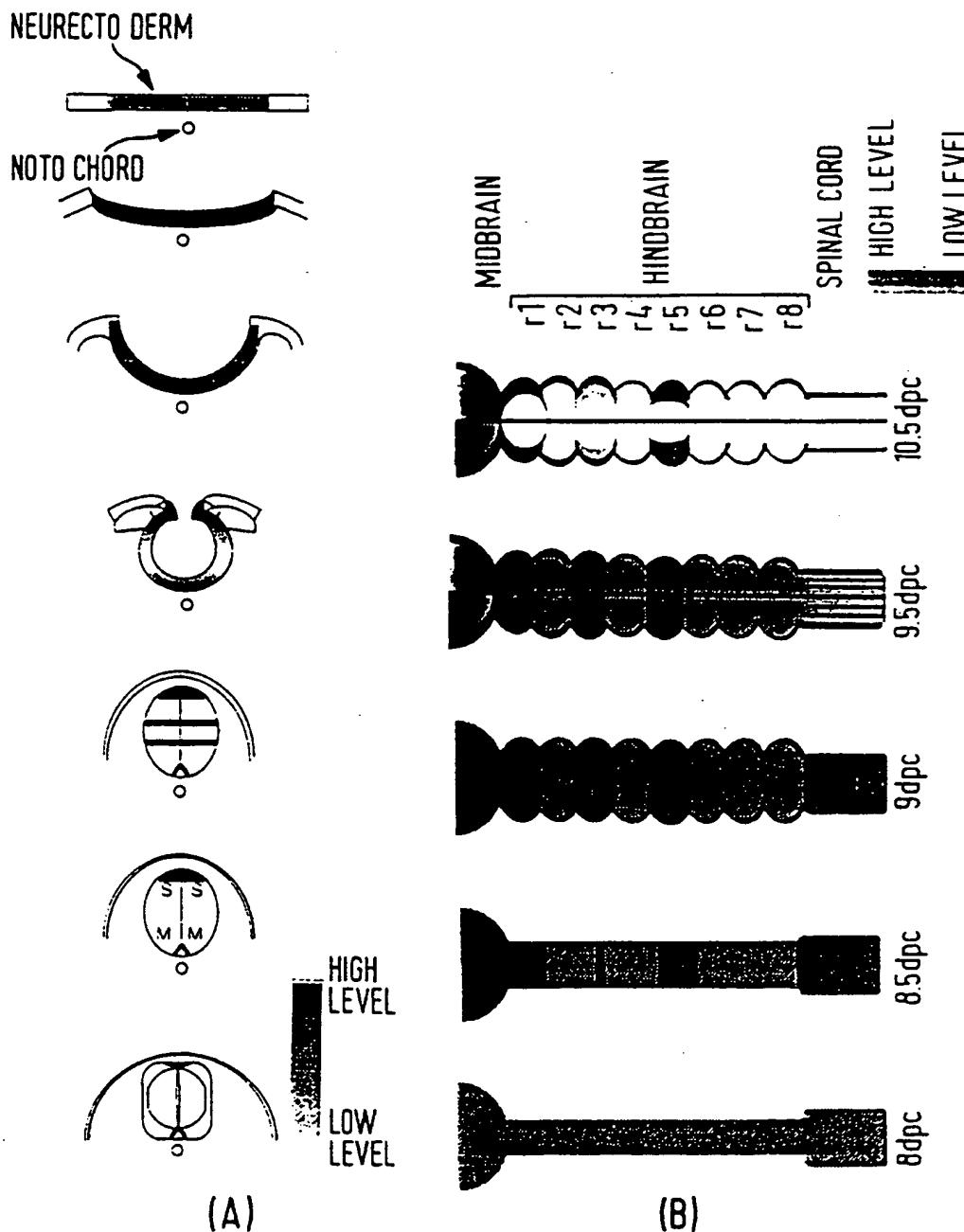
FIG. 1

FIG. 3



3 / 17

FIG. 4

EXPRESSION OF ME2 IN THE DEVELOPING MOUSE
SPINAL CORD (A) AND HINDBRAIN (B)

4/17

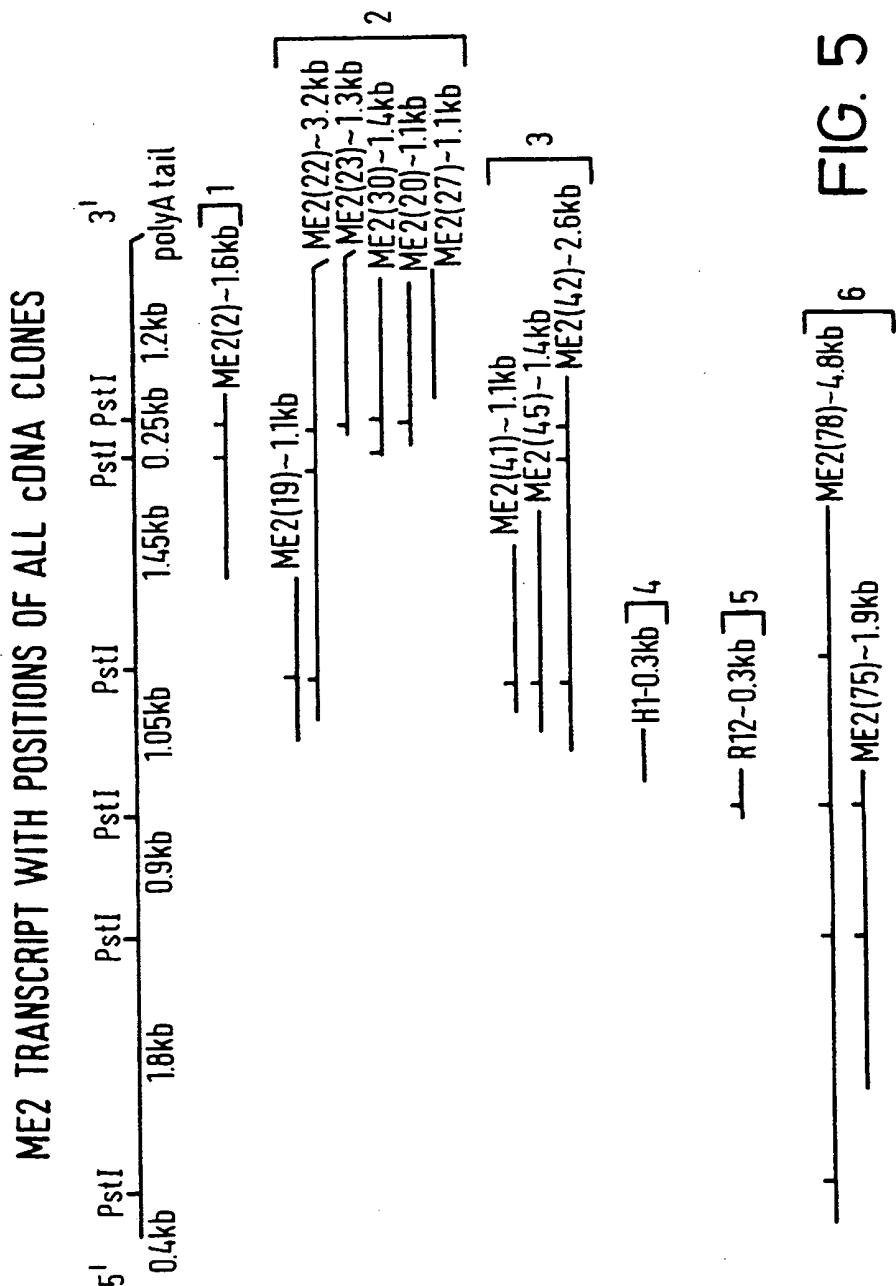


FIG. 5

5/17

AMINO ACID SEQUENCE OF ME2 REGION COVERED BY ME2(78) AND ME2(22)

1573 RESIDUES

GDYCETEIDLCSNPCGANGGCRSREGGYTCECFEDFTGEHCQVNVRSGRCASGVCK
NGGTCVNLLIGGFHCVCPPGEYEHPYCEVSTRSFPPQSFTFRGLRORFHFTVSLAF
ATQDRNALLLYNGRFNEKDFIALEIVEEQLQLTFSAGETTTVTPOVPGGVSDGRW
HSVLVQYYNKPNIHLGLPHGPSGEKVAVTVDDCDAAVAVHFGSYVGNYSCAAQGT
QSGSKKSLDLTGPLLGGVPNLPEDFPVHSRQFVGCMRNLSIDGRIVDMAIFIANNG
TRAGCASQRNFCDGTCQNGGTCVNRWNTYLCECPLRFGGKNCEQAMPHPORFTGES
VVLWSQDLDITISVPWYLGLMFRTKEDGVLMETAGTSSRLHLQILNSYIRFEVSYG
PSDVASMQLSKSRITDGGWHHLIELRSAKEGKDIKYLAVMTLDYGMDOSTVQIGNQ
LPGLKMRTIVIGGVTEDKVSVRHGFRGCMQGVRMGESSTNIATLNMMNDALKVRVKDG
CDVEDPCASSPCPPHRPCRDTWDYSICDRGYLEKKCVDACLLNPCKHVGSLCALP
NTPRGYSCCECPGPHGQYCENKVDLPCPKGWWGNRCVAPVTVLAKALIPTATRPMA
SARRITTSPOPRIVAFPTVSPRSHSRACDMDTGQCACKPGVIGROCMRCDNPFAEV
TSLGCEVIYNGCPRAFEAGIWWPQMKFGQPAAVLCPKGSGVNAVRHCSGEKGWLPP
LFNCTSGSFVDLKALNEKLNNETRMDGNRSLRALKRNATQGNSTLFGNDRVTAY
QLLARILOHESRQOGFDLAATREANFHEDVVHTGSALLAPATEASWEQIQRSEAGAA
QLLRHFEAYFSNVARNVKRTYLRPFVIVTANMILAVDIFDKLNFTGAQVPRFEDIQE
ELPRELESSVSPADTFKPPEKKEGPVVRLTNRRTPLTQOPEPRAERETSSSSRRR
HPDEPGQFAVALVVIYRTLGQLLPEHYDPDHRSLRLPNRPVINTPVVSAMVYSEGTP
LPSSLORPILVEFSLLETEERSKPVCVFWNHSLDTGGTGGWSAKGCELLSRNRTHVT
COCSHSASCALMDISRREHGEVPLPLKIITYAALSLSLVALLVAFVLLSLVRTLRSN
LHSIPOEPIHALFFSOLIFMVGINOTENPFLCTVVAILLHYVSMGTFAWTLVENLHV
YRMLTEVRNIDTGPMAFYHVGWGIPIAVTGLAVGLDPQGYGNPDFCWLSQLQDTLIW
SFAGPVGTVIIINTVIFVLSAKVSCQRKHYYERKGVVSMLRTAFLLLLVTATWLL
GLLAVNSDTLSFHYLFAAFSCLOGIFVLLFYCVANREVRKHLRAVLAGKKLQLDDSA
TTRATLLTRSLNCNNTYSEGSRHAPHRPGQSTASLDSTRDEGVQKLSVSSGPARGN
HGEPDASFIPRNSKKAHGPDSDSDESLSLDEHSSSSYASSHTSDSEDDGGEAEDKWNP
AGGPAHSTPKADALANHVPAGWPDESLAGSDSEELDTEPHLKVRPRSAWSYTGRRA
ITVATGPLTRKVGSWPSQWPCLAASPRSSGKAS*

SEQ ID NO. 1

FIGURE 6

6/17

COMPLETE NUCLEIC ACID SEQUENCE OF ME2 REGION COVERED BY ME2(78) AND ME2(22) cDNA CLONES

NUCLEOTIDES 1 - 6791

CCGGGGACTACTGCGAGACTGAAATTGATCTTGCTACTCCAATCCGTGCGGGGCCA
ATGGCGCTGCCGGAGCGTGAAGGTGGCTACACTTGAGTGCTCGAGGACTTCA
CTGGGGAGCATTGCCAGGTGAACGTTGCTCAGGCCCTGTGCCAGCGAGTATGCA
AAACGGGGCACCTCGTGAACCTGCTATTGGAGGCTTCCACTGTGTGCCCCGC
CCGGCAGAGTATGAGCATCCCTACTGTGAAGTGAGCACCAAGGAGCTCCACCGT
CCTCGTTACCTTCCGAGGCCCTGCCGCAACGCTTCACTTACCGTCTCCCTGGC
TTGCCACCCAGGACAGGAATGCCCTGCTGCTCTACAATGGCCGCTTCAATGAGAAGC
ACGACTCATGCCCTGGAGATTGGAGGAGCAGCTGCAGCTACGTTCTGGCAG
GTGAGACCAACCAACCGTGAACCGCAGGTTCTGGAGGTTGTGAGCGATGGGCGG
GGCATTCCGGTGGTGCAGTACTACAACAGCCAAACATTGGCACCTGGCCTGC
CCCACGGGGCGTCTGGAGAGAAGGGTGGTGTGGTACTGTGGATGACTGTGACCGAG
CGGTGGCGTGCACTTGGAGTTACGTGGGAACTACAGCTGCCGTGCCAGGGCA
CTCAGAGCGCCTCCAAGAAGTCACTGGATCTGACTGGCTCTGCTCTGGTGGT
TCCCCAACCTGCCAGAAGACTTCCCGTGCACAGCGTCAGTTGTGGGATGCATGC
GAAACCTGTCATCGTGGCCGGATTGTGGACATGGCTGCCTTATTGCCAACATG
GTACCAAGGGCAGGTGTGCTTCTCAAGGAACCTCTGCATGGGACCTCATGCCAGA
ACGGGGGACCTGTGTGAACAGTGAAACAGTCACTTATGTGAGTGCCCGTCCG
TTGGTGGAAAAGAACTGTGAACAAAGCTATGCCAACACCTCAGCGCTTCACTGGT
GCGTCGTGTTGGAGGTGGACCTTGACATCACCATCTGTGCCTTGGTACCTGGG
TCATGTTCCGGACCCGGAGGGAGGATGGTGTGCTATGGAAGGCCACAGCTGGCAC
CTTCCAGGCTCCATCTCCAGATTCTCAACAGCTACATCCGCTTGGAGGTCTCTAC
GCCCTCTGACGTGGCATCCATCGAGCTGTCCAAGTCCGGATAACTGACGGGG
GGCATCACCTGCTATAGAACTGAGGGAGTGCCAAGGGAGGCAAGGACATCAAATACC
TGGCAGTCATGACCTGGACTATGGATGGACAGAGCACAGTGCAGATTGGGAAATC
AGCTTCCGGGGTGAAGATGGGACATTGGTATGCAAGGGAGTGTGACCGAGGAAAG
TCTCTGTCGCCATGGTTCCGAGGCTGTATGCAAGGGAGTGTGAGGATGGGAGAGAG
CCACCAACATTGCCACCCCTGAACATGAATGACGCCCTCAAGGTCAAGGTGAAGGAG
GCTGTGATGGAGGACCCATGTGCTCAAGCCCTGCCCTCCCCATAGACCTGCC
GTGACACATGGGACAGCTACTCTGCATCTGACAGAGGGTACTTGGAAAAAAAGT
GTGTTGGATGGCTGTCTCTGAACCCCTGAAGCACGTCAGTGGGAGCTGTGCGCTCC
CCAACACTCTGAGGCTACTCTGGAGTGGGACCCGGCAACTATGGGAGTACT
GTGAGAGACAAAGTCGACCTTCCGTGCCCCAAAGGTGGTGGGGAAACCGGTGTGG
CCCTGTCACTGTGCTGCAAGGCTTGTGATCCGACTGCAACAAAGACCAATGG
CCAGTGCAGGAGAATTACTACAAGCCCCAGCCAGGATCGTTGCCCTGGTGA
CTGTTCCCCCGCTCCACAGCGTGCCTGCACTGGACACTGGCAGTGTGCC
GCAAGCCTGGTGTATGGCCGTCACTGGCAACCGCTGTGATAATCTTCTGGGAGG
TCACCTCGCTGGCTGTGAAGTGTACTACAATGGGTGCTCCAGAGCATTTGAGGCTG
GCATCTGGTGGCCACAGATGAAATTGGCAGCCAGCAGCGGTCTATGCCCAAG
GATCCGTGGTAAACGCACTGGCACTGCAAGTGGGAGAAGGGCTGGCTTCCCCAG
AGCTTCAACTGCACCTCTGGCTTTGTGGACCTCAAGGCCCTGAACGAGAAC
TGAACCGCAACGAGACAAAGAATGGACGGGAACCGGCTCTGCGCTGGCAAAGGCTC
TGAGGAACGCCACGCAAGGGAAACAGCACCCCTTTGGCAATGATGTGCGCACGGCCT
ACAGCTTCTGGCCCGCATCTTACAGCATGAGAGCCAGCAGGGCTTGCACCTGG
CAGGCCACGGAGGGCTAATTTCATGAGGATGTGCTCATACAGGCAGCGCCCTCC
TGGCCCCAGCTACAGAGGCACTGTTGGGAACAGATCCAGCGAAGGGAGGCTGTG
CGCAGCTACTGAGGCATTGAGGCATACTTCAGCAACGTTGGCACGAAATGTGAAGA
GGACCTATCTGAGGCCCTTGTATCGTACCGCCAACATGATTGAGTTGAGGACATTGAG
TCTTCAGACAAGCTGAACCTCACGGGTGCCAGGTGCCAAGGTTTGGAGGACATTGAG
AAGAGCTCCAAGGGAGCTGGAGTCTCCGTGCTTCCAGCTGACACCTTCAAGC

SEQ ID NO 2

FIGURE 7

CAACAGAGAAAAAGAAGGCCCGTGGTGAGGCTGACCAACCGGAGGACTACCCAC
TCACGGCACACCAGAGCCCAGGGCTGAGAGGGAAACCTCATCAGCAGACGGAGGA
GACACCCCGATGAGCCTGGACAGTTGCTGTTGCCCTGGTTGTCATTTACCGGACCC
TGGGTCACTGCTGCCATGAAACTATGACCCCGACCATCGCAGCCTCCGACTGCCTA
ACCGGCCCTGTCATCAACACCCCGTGGTGAGTGCATGGTGACAGTGAGGGAAACCC
CACTGGCCAGCTCTGCAAGGGCTATCCGGTGAGTTCTCCCTGGAGACGG
AGGAACGAAAGCAACCTCTGTCATGGAAACCACTCCCTGACACTGGTGAGGA
CTGGAGGGTGGTCAAGGCCAGGGCTGAGTCTGAGGAAACGGGACCCACGTCA
CTTGCCAGTCAGCATTGCCAGCTGCGCGTGCTCATGGACATTTCCAGACGTCA
AGCACGGGGAGGTTCTGCCCTGAAGATCATCACCTATGCCGCCCTGTCCTTGCTT
TGGTGGCCCTCTGGTGGCTTGTCTCTCGCTCGTCACTGCCACTGCGCTTCA
ACCTGCACAGCATCCCACAAGAACCTATCCACGCTCTGTTCTTCCAGCTCATCT
TCATGGTGGCATCACCCAGACTGAGAACCCGTTCTGCACTGGTGCCTGCCATCC
TCCTGCACTACGCTCATGGCACCTCGCCTGGACCCCTGTTGGAGAACTTGATG
TCTACCGCATGTCAGAAGTGCGCAACATCGACACTGGGCCATGGCGTTCTACCC
ACGTGGTGGGCTGGGGCATCTGCCATTGTCAACAGGACTGGCTTGGGCTTGGGACCC
CTCAGGGCTATGGAAACCCCTGACTTCTGCTGGCTGCCCCCTCAGGATACCCCTGATT
GGAGCTTGCTGGGCTGCGAACGGTTATAATCATCAACACAGTCATTTGTCC
TGTCTGCAAAGGTTCTGCCAAAGAACGACCAATTATGAAAGAAAGGGGGTTG
TCTCCATGCTGAGGACGGCCTTCTCTGCTGCTGCTCGTCACTGCCACCTGGCTGC
TGGGACTGCTGGCGGTCACAGTGAACACTCTTAGCTTCACTACCTCTTGCTGCCT
TCAGCTGCTTGAGGGCATCTTGCTCTCTGTTACTGCGTGCCAAACAGGGAGG
TGCAGGAAACCTGAGGGCGGTGCTGGCAGGGAAAGCTGCACTGGATGACTCGG
CCACACTCGGGCCACTCTGTAACCGCCTCCCTCAACTGCAACAAACACCTACAGCG
AAGGGTCAGACATGCTCCGACCGCCCTGGGCACTGGCACAGCTCTGAGAAGTA
CCACCAAGGGATGAAGGGGTCCAGAAACTCAGTGTGCTCTGGCCAGCCGGTGGTA
ACCATGGAGAACAGATGCACTCTCATCCCTAGGAACCTCAAAAGCTCACGGCC
CTGACTCTGACTCTGACAGTGAGCTGCCCCGGACAGCAGTAGTTCTACGGCCT
CTTCACACACATCGGACAGCGAGGATGATGGCGGGAGGGCTGAAGACAAATGGAATC
CGGCTGGGGCCCCCCTAGCACCCAAAAGCAGATGCTGGCCAACCCAGTC
CAGCTGGCTGGGGCACAGAGAGCCTGGCTGGAGTGAAGCTGAGGAGTTGGACACTG
AGCCCCACCTGAAGGTAGGCAAGGCTAGCGTGAGGTTACACGGCAGGGCAGGG
CAATCACTGTGGGAGCCGGCTCTGACCCGGAAAGTGGGGCTTGGCCAAGCAGT
GGCGTGGCTTAGCAGCCAGCCCCAGGGAGCAGCGGAAAGGATCTGAAAAACAAAGT
CACCTACCCGCCGGCATTGCCAGAGCAGCCACTGAAGTCCGGCTGCGAGAGAAC
TGGCTGATTGTAGCAGAGCCCCACATCCTCCGACATCCTCCCTGGCTTGGG
ATGGTGCTCATGCCACTGACTGTGCTTACCATCAAGACTCCGAGGGAGGGAGCCAG
GCCGTGAGCATCTCAATGGGGTGGCATGAATGTAACGCACAGGGAGTGGCAGGGCAA
CGGTTCTGACTGAGAACCATGAGGGCACAGTCACCCCAACAGACTGCCGGTCAAG
CCCTCAGACCTGAAGGGCTGCCACTGGACTGCTGCCTATGGGACAGCAGGCCACTTG
TGTGAGGGTCCCTGCCATAGCAGCTGGCTCTACGGCAGACGGTCCAGCAGGGAAAGCCC
TTGACCCCTCATAGGAGCTAGGGCCAGACTCTGACAAAGTGCCAAAGCCACAGATG
TCTCAGAGGGAGACGTGGACTTCAAGGCTGGCATAGCTCGTCCCTTAGACTGA
AGACAGAAATCAAACCATGCTGACAGGGCATTGAGCAGAGCTGGACTTGGT
AATCATTGTACCGGGCCCTCAACTGTCGGCAGGGCTCTCTTGTGGTACAAGCCC
ATCACCACCGCAGCTAGGGTGGCTCTGCAACGGCAACCCCTGGGTTAAATTGCTGT
CTAAAAATGTAATAGATAAAATCTCCCTGGACTTGGGAGAAGATGGGAGCTGT
GTATGCTTACACTGCTTACTGCACTGGAGAGGCCATGAAATGGCATCTC
ACTCTATTGCCAAGGAAGCTGACAGTTGACTTGAATCTGGAAATGAGTCACACT
AGCTGGTCCCAGTGCCAGGTAGGGAGCATGGGCTGTGAAGTTGACACAGCTTCT
GCAGCTGCTCCGGCAGGGAGTGTGGGTGTTCCGCTGGCTTGGAGACACGGCA
CATGTGCTCGTATGTGTGCCAGCTCCATGGATCCAGGTCAAGGGCTCACCTGTGA
GGAGTGGGGCCATAGCTATGATAAGAACCTGACCAACCCCTGCCACCCCCACGGCC
CCCCCCCCCCCCGCTCTGAGTGCAGTGAGGGCAGTCTGGAGCCTTCCCAGTCAGC

FIGURE 7 CONTINUED

CCAGTGCGCAGCGGGAGGACTTCTGCTCCTGCTCCATAAGCTCTCAGAGTGCTCAT
GTTCCCTGATTCTGAGGGAGGCCAGAGGTCTTCCAAACGCTGTCAGGGCTTCTGCC
AATGTTGACTTTTCAAGACTCCAGGGCTCTCTAGGCAGGCAGGTGCTCACCCTG
AAATCTGACCAATGACATCACTTGTCTCAAATGACCAATTGTGCAAAGAAACAAAG
CCAGAGTGCCGTTTCAATGGTTACACATTCTTTGGAAATCTCACACCAAGGAC
CTGTGACCCAGCCACTGAGAGGCCACGGTGAGCCAAGGGCAAGGGATGGGAGCCTGGA
GTTACTCAGACAGCTTACTTCAGCATGGACTGTTGCTGAATCAGGTCCCCAAAGT
ACATGGGTGCACAGTCGCTCGGAATGGAGAACCTCAGGCAGGGCGGTCAAAGGCCA
GGACTTCTACCAAGCTGTGCTCTCAGAAGTGACAGGACTGCTAGTGACTGACTGG
TGAGATGAAAGCTACAAGACATGGCTGGGGTCACGCACTGCCAAGAGGCTGACGG
GAGGCCAGGTAGCCAAAGGATGGCAAGGATAAGAGTGACCTAGCACAGGGGAGCT
TCAGTCCCAGGTGGTACAGCACCGTGACAACCTCCGCAACCCACGCCACCTCAGAAGG
TGAAGTTTGTGATCACAACATTAGCAAACAAACCCCTGTCAGTTTAAACT
TTTATTACTGTTGACGCACTGGATGGCAACACAGGTGAGATGATGACATCTATAATAAA
TTAAGATTTGGATTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAA

FIGURE 7 CONTINUED

9/17

ME2 PROTEIN 1-1798 AMINO ACIDS

NSGLMEKLQIEEQTKKAQQELEEQTRRALELEQERKAGTAIELRAHDPDEGDAGRLSYQMEALFDER
SNGYFLIDAATGAVTTARSLDRETKDTHVLKVSADHGSPRRSAATYLTVTSNDHSPVFEQSEYRER
IRENLEVGYEVLTIRATDGAPSANMRYRLLEGAGGVFEIDARSGVVRTRAVVDREAAEYQLLVEAND
QGRNPGPLSASATVHIVVEDENDNPQFSEKRYVQVPEDAVNTAVLRVQATDRDQGQNAAIHYSIVSG
NLKGOFYHSLSGSLDVINPLDFEAIREYTLRIKAQDGGRPPLINSSGLVSQVLDVNDNAPIFVSSPFQ
AAVLENVPLGHGSVLHIQAVDADAGENARLQYRLVDTASTIVGGSSVDSENAPASADFPFQIHNSSGWITV
CAELDREEVEHYSFGVEAVDHGSPAMSSASVSITVLVDVNDNDPMFTQPVYELRLNEDAAGSSVLT
RDRDANSVITYQLTGGNTRNRALSSOSGGGLITLALPLDYKQEROYVLA
VTSADGTRSHTAQVFINVTD
ANTHRPVFOSSHVTVSSEDGPVGTISIATISATDEDTGENARITYVLED
PVPQFRIDPDTGTIYTMTEL
YEDQAAYTLAITAQDNGIPOKSDDTSLEI
LILDANDNAPRFLRDFYQGSVFEDAPPSTSVLQVSATDRDS
GPNRGLLYTFQGGDDGDFYIEPTSGVIRTO
RRLDREN
VAVYNLWALAVDRGSPNPLSASVGQIVSVLD
INDNPPVFEKDELELFVEENSPVGSSVARIRANDPDEGPNAQIYQIVE
GNVPEV
FQLDLSGDLRALVE
LDFEVRRDYM
L
VQATSA
PLV
SRATV
H
IRL
L
D
Q
N
P
E
L
D
P
A
T
G
E
L
Q
N
D
N
P
E
L
D
R
E
N
V
A
V
Y
N
K
P
N
I
G
H
L
G
L
P
H
G
P
S
G
E
K
V
A
V
V
T
D
D
C
A
A
V
V
H
F
G
S
Y
V
G
N
Y
S
C
A
A
Q
G
T
Q
S
G
S
K
S
L
D
T
G
P
L
L
G
V
P
N
L
P
E
D
F
P
V
H
S
R
Q
F
V
G
C
M
R
I
D
G
R
I
V
D
M
A
F
I
A
N
G
T
R
A
G
C
A
S
Q
R
N
F
C
D
G
T
S
C
O
N
G
G
T
C
V
N
R
W
N
T
Y
L
C
E
C
P
L
R
F
G
G
K
N
C
E
Q
A
M
P
H
P
R
F
G
E
S
V
V
L
W
S
L
D
I
T
I
S
V
P
W
Y
L
G
M
F
R
T
R
K
E
D
G
V
L
M
E
A
T
G
T
S
S
R
L
H
L
Q
I
L
N
S
Y
I
R
F
E
V
S
Y
G
P
S
D
V
A
S
M
L
S
K
S
R
I
T
D
G
G
W
H
H
L
I
E
L
R
S
A
E
G
K
D
I
K
Y
L
A
V
M
T
L
D
Y
G
M
D
Q
S
T
V
Q
I
G
N
Q
L
P
G
L
K
M
R
T
I
V
I
G
G
G
T
E
D
K
V
S
R
H
F
R
G
C
M
Q
G
V
R
M
G
E
S
T
N
I
A
L
N
M
N
D
A
L
K
V
R
V
K
D
G
C
D
V
E
D
P
C
A
S
P
C
P
H
R
P
C
R
D
T
W
D
S
Y
S
C
I
C
D
R
G
Y
L
E
K
K
C
V
D
A
C
L
L
N
P
C
K
H
G
S
L
C
A
L
P
N
T
P
R
G
Y
S
C
E
C
G
P
H
Y
G
Q
Y
C
E
N
K
V
D
L
P
C
P
K
G
W
G
N
R
C
V
A
P
V
T
L
S
A
K
I
P
T
A
T
R
P
M
A
S
R
I
T
T
S
P
Q
P
R
I
V
A
F
P
V
T
V
S
P
R
S
H
R
A
C
D
M
T
G
Q
C
A
C
K
P
G
V
I
G
R
O
C
N
R
C
D
N
P
F
A
E
V
T
S
L
G
C
E
V
I
Y
N
G
C
P
R
A
F
E
A
G
I
W
W
P
Q
M
K
F
Q
P
A
A
V
L
C
P
K
G
S
V
G
N
A
V
R
H
C
S
G
E
K
G
W
L
P
P
E
L
F
N
C
T
S
G
F
V
D
L
K
A
L
R
A
N
T
Q
G
N
S
T
L
F
G
N
D
V
R
T
A
Y
Q
L
A
R
I
L
Q
H
E
S
R
Q
Q
G
F
D
L
A
A
T
R
E
A
N
F
H
E
D
V
V
H
T
G
S
A
L
L
A
P
A
T
E
A
S
W
E
Q
I
Q
R
S
E
A
G
A
A
Q
L
R
H
F
E
A
Y
F
S
N
V
A
R
N
V
K
R
T
Y
L
R
P
F
V
I
V
T
A
N
M
I
L
A
V
D
I
F
D
K
L
N
F
G
Q
A
V
Q
P
R
F
E
D
I
O
E
E
L
P
R
E
L
E
S
S
V
F
P
A
D
T
F
K
P
K
E
G
P
V
V
R
L
T
N
R
R
T
P
L
T
A
Q
P
E
P
R
A
E
R
E
T
S
S
S
R
H
P
G
O
S
T
A
L
D
S
T
T
R
D
E
G
V
Q
K
L
S
V
S
G
P
A
R
G
N
H
G
E
P
D
A
S
F
I
P
R
N
S
K
K
A
H
G
P
D
S
D
S
E
L
S
D
E
H
S
S
S
Y
A
S
S
H
T
S
D
E
D
G
G
E
A
E
D
K
W
N
P
A
G
G
P
A
H
S
T
P
K
A
D
A
L
H
V
P
A
G
W
P
D
E
L
A
G
S
D
S
E
L
D
T
E
P
H
L
K
V
R
G
S
W
P
S
Q
P
C
L
A
A
S
P
R
S
G
K
A
S
*

SEQ ID NO. 3

FIGURE 8

ME2 DNA SEQUENCE 1 - 8210

GAATTCCGGGCTGATGGAGAAGCTGAAGCAGATTGAGGGAGCAGACTAAGAAGGCTCAGC
 AAGAGCTGGAAGAGCAGACCCGAGGGCCCTAGAACCTTGAGCAGGAACGGAGCGTGGC
 GGCACCTGCGGTACATCGAACCTGCGCGCAGCACCCAGAGCAAGGCGATGCAGGACGCC
 TAGCTACCAAGATGGAGGCCTGTTGATGAGCGCTCAATGGCTACTTCCTCATCGATG
 CGGCCACGGGCTGAGTACAGCACAGCCCGCTCCCTGGACCCGGAGACCAAGGACACTCAT
 GTACTCAAAGTTAGTGTGAGCACGGCTCCCGAGGGCTCAGCTGCCACCTACCT
 CACCGTAACACTGTCAGTACACTAACGACCAACAGGCCAGTCTTGAGCAGTCTGAGTATC
 GAGAGCGAATTGAGAAACCTGGAGGTGGGCTATGAGTTCTGACCATCGTGGCCACC
 GACGGGGATGCCCTTCAACGAAACATGCGTATCGTCTGCTGGAGGGCGCAGGTGG
 TGTCTTGAGATAGACGCACGATCAGGTGTCGTGCGCACACGGGCTGTGGTGGACCGTG
 AGGAGGCGGCTGAGTACAGCTGCTGGTGGAGGCCATGACCAAGGGTGCACATCAGGC
 CCACTCACTGCTCACGGCACCGTCCACATAGTGGTAGAAAGCAGAGAATGACAACCTACCC
 CCAGTTCACTGAGAAGGCTATGTTCAAGTCCAGAAGACGTAGCCGTCAACACGG
 CTGTGCTTCACTGAGGCTCAGGCAACTGACCCGGACAGGGCAGAATGCAGCCATACACTAC
 AGCATCGTTAGTGGCAACCTGAAGGGCTAGTTCACCTGCATTGCTTAGTGGAGCCT
 GGATGTTATCAACCCGCTGGACTTCAAGCCATCCGGAAATACACCCCTGCGCATCAAAG
 CCCAAGATGGGGCCGCGCTCCTCTTCAATTCTCAGGACTGGTCTGGTGCAGGTG
 TTAGATGTGAACGACAATGCGCCATTTGTGAGCAGCCCTTCACTGGCTGCCGTGCT
 AGAGAATGTGCCCCCTGGCCACTCAGTCTGACATCCAAGCAGGTGGACGAGATGAG
 GGGAGAACGCCAGGCTCAGTACCGCTAGTGGACACAGCCTCCACTATGTTGGGGG
 AGCAGTGTGCACTGTGAGAACCTGCGCTCTGCCCCAGACTTCCCAATCACAA
 CAGCTCCGGTTGGATTACTGTGTCGGGAGCTGGACCGTGGAGGGTGGAAACACTATA
 GCTTGGAGTAGAAGCACTGGACCATGGCTACAGGCATGAGCTCTCTGCCAGCGTG
 TCCATCACAGTGTGGATGAAATGATAACGACCCCATGTTCACGCGACCTGTGATGA
 GCTGCGCTGTAATGAGGATGCGGCTGCGGAGCAGCGTGTGACCCCTCAGGGCCGAG
 ACCGTGATGCCAATAGTGTGATCACCTACCGACTGACGGGTGGGAAACACCCGCAACCGC
 TTGCACTCAGCAGCACAGCGCCGTTATCACCTTGGCACTGCCCTGGACTA
 CAAGCAGGAACGGCAGTATGTGTCGGCTGTGACCCGCTGGACGCGTGTACACA
 CCGCGCAGGTCTTATCACAGTGGCAACACCCACAGGCCGGTTTCCAGAGT
 TCCCACTACCGTCACTGTGAGTGAAGACGCCGGCTGGCACCTCATCGTACCAT
 CAGTGCACGGATGAGGATAACGGGTGAGAACGCCGCACTCACCTATGTGAGGATC
 CCGTACACAGTCCGATTGACCCGGACACTGGCACCAATTACCATGACGGAACTG
 GACTATGAGGACAGGGCTGCTACACGCTGGCCATCACGGCTCAGGACAATGGCATTCC
 TCAGAAGTCAGACACTACCTCTCTGGAGATCTTACCTCGACGCCAATGACAACGCC
 CCAGGTTCTGGAGATTCTACCGGGGTTCTGTTTGAGGATGCCGGCTTACACCTTGGG
 AGTGTCTCCAGGGTCTGCTACAGACGGTGAATGGCCCTAATGGCCGCTCTGTG
 CACTTCCAGGGTGGGAATGAGGATTTCTACATTGAGCCACGTCCTGGTG
 TGATCCGTAACCCAGCGCCGGCTGGACAGAGAGAATGTGGCGTGTACACCTTGGG
 CTCGCTGTTGAGATGGGGAGCCGAATCCCTCAGTGCCTGAGTGGGAAATTAGGTGAG
 TGTGTTGGACATTAACGACAACCCCCCAGTGGAGAAGGACGAGCTGGAGCTGTTG
 TGGAGAGAACAGCCCTGGGTTAGTGGTAGCAAGAATAGGGCAACGCCGGAC
 GAAGGTCGAATGCTCAGATCATTATCAGATGCTGGAGGGCAATGTGCCCAGGGTCTT
 CCAGCTGGACCTACTGAGTGGTGACCTCGCTGCCCTGGCTGAGTTGGAGGTT
 GGAGGACTATATGTTGGTGTGAGGCCACGCTGCTCTCTGGTAAGCCGGCCACC
 GTGACATCCGCTCTGGACCAAGAATGACAACCCACCGAGTTGCCGACTTCCAGAT
 CCTTTCACAACACTATGTACCAATAATCCAACAGCTTCCCACTGGTGTGATCGGC
 GCATCCAGCCACGACCCCTGACCTATGACAGCTCAATTACACCTTCTGCAAGGC
 AACGAGCTGAGCTGCTGCTGCTGAGTGGATCCGCCACAGGAGAGTTGAGCTCAGCCGG
 TCTGGACAAACACGGGCACTGGAGGCGCTCATGGAGGTGTCTGTCAGATGGTATCC
 ACAGCGTACCGCTCTGCACTCTGGCGTGCACATCAGGAGATGACATGTC
 AACAGCATCACTGTCCGCCCTGGAGAACATGTCGAGGAGAACGTTCTGCTCCCGCTG
 GTCCCTCTTGTAGAAGGGGTGGCCACAGTACTGTCACCCACCAAGGAGTGCACATCTCG
 TCTTCAACATCCAGAACGACACGGACGCTCAGCTCAACATCTGAAACGTGACTTCTCG
 GCACTGCTCCCGGTGGCACCCGTGGCCGGTTCTCCGCTGAGAGGACCTGCAAGGAGCA
 GATCTACCTGAACCCGGACACTGCTCACCCACATCTCCGCCAGCGTGTGCTGCCCTTG
 ATGACAACATCTGCGTGGAGGGAGCCCTGCGAGAACTACATGAAGTGCCTGCTCGT

SEQ IN NO.4

FIG 9

AGGTTTACAGTTCCGCACCCCTCATTAGTCCACCGGTGCTTCCGGCTATCCA
 TCCCACACGGGCCTGCCTGCCCTGCCGGGTTTACCGGGGACTACTGCGAGA
 CTGAATGGATCTTGTACTCCAATCCGTGCGGGCAATGGGGCTGCCAGCGT
 GAGGGTGGCTACACTTGTGAGTGCTCGAGGACTTCACTGGGGAGCATTGCCAGGTGAA
 CGTCGCTCAGGCCGCTGTGCCAGCGGAGTATGCAAAACGGGGCACCTGCGTGAACC
 TGCTATTGGAGGCCTTCACTGTGTGCCCCGGCGAGTATGAGCATCCCTACTGT
 GAAGTGAGCACCCAGGAGCTTCCACCCAGTCCCTGTTACCTCCGAGGCCCTGCC
 ACGCTTCACTCACCGTCTCCCTGCCGTTGCCACCCAGGACAGGAATGCCCTGCTGC
 TCTACAAATGCCGCTCAATGAGAACGACACTTCATGCCCTGGAGATTGTGGAGGAG
 CAGCTCAGCTCACGTTCTCGGAGGTAGAGACCAACCCAGGTGACACCCGAGGTTCC
 TGGAGGGTGTAGCGATGGCGGTGGCATTGGTGTGGTGCAGTACTACAACAAGCCA
 ACATTGGCCACCTGGGCTGCCACGGGGCTCTGGAGAGAAGGTGGTGTGGTACT
 GTGGATGACTGTGACGCCAGCGTGGCGTGACTTGGAGTTACGTGGGAACTACAG
 CTGCCGCTGCCAGGGCACTCAGAGCGGTCCAAGAAGTCACTGGATCTGACTGGTCTC
 TGCTTCTGGTGTGGTCCCCAACCTGCCAGAACGACTTCCCGTGACACCCGTCAGTT
 GTGGGATGCGAACCTGGTCCATCGATGCCGATTGTGGACATGGCTGCGTTTAT
 TGCAACAAATGGTACCGCAGGCTGTCTCAGAGGAACCTTGCGATGGGACCT
 CATGCCAGAACGGGGCACCTGTGTAACAGGTGAAACACGACTTATGTGAGTGCCG
 CTCCGCTTGGTGGAAAGAACGTGTAACAGCTATGCCACACCCCTCAGGCTTCACTGG
 TGAGAGCGTCGTGGTGTGGAGTGACCTTGACATCACCATTCTGTGCTTGGTACCTGG
 GGCTCATGTTCCGACCCGGAAGGAGGATGGTGTGTGATGGAAGGCCACAGCTGGCAGC
 TCTCCAGGCTCATCTCCAGATTCTCAACAGCTACATCCGCTTGGTCTCTACGG
 CCCCTGACGTGGCATCCATGCCAGCTGTCAAGTCCCGGATAACTGACGGGGGTGCG
 ATCACCTGCTAGAACACTGAGGACTGCCAAGGGGCAAGGACATCAAATACCTGGCA
 GTCATGACCTGGACTATGGGATGGACAGGACAGTGCAGATTGGGAACTAGCTTCC
 TGGGTTGAAGATGCGGACTATTGTGATCGGAGGTGTGACCGAGGACAAGGTCTGTCC
 GCCATGGTTCCGAGGCTGTATGCAAGGGAGTGGAGAGAGAGCTCCACCAACATT
 GCCACCCCTGAACATGAATGACGCCCTCAAGGTCAAGGTGAAGGACCGCTGTGATGTGGA
 GGACCCATGTGCTCAAGCCCTGCCCTCCCATAGACCCCTGCCGTGACACATGGGACA
 GCTACTCTGCATCTGTGACAGAGGGTACTTGGAAAAAAAGTGTGTGGATGCGTGTCTC
 CTGAACCCCTGCAAGCACGTTGGCAGCTGTGCTGCCCTGCTCCCAACACTCCCTGAGGCTA
 CTCCGAGCTGGGACCCGCCACTATGGGAGTACTGTGAGAACAAAGTCGACCC
 CGTCCCCCAAAGGCTGTGGGGAAAGGTGTGCCCCCTGTCACTGTGCTGTGAGCC
 AAGGCTTGTCCCAGCTGCAACAGACCAATGCCAGTGCAAGGAGAAATTACTACAAG
 CCCCCAGCCCAGGATCGTGCCTCCCTGTGACTGTTCCCCCGCTCCACAGCCGTG
 CCTGCAACATGGACACTGGCAGTGTGCTGCAAGGCTGGTGTCACTGGCGTCACTG
 AACCGCTGTGATAATCTTCGCGGAGGTCACTCGCTCGGCTGTGAAGTGAATCACAA
 TGGGTGTCCCAGAGCAATTGAGGTGGCATCTGGTGGCACAGATGAAATTGGGAGC
 CAGCAGGGTGTATGCCCAAAGGATGGTGGTAACGCACTGCCGACTTGCACTGG
 GAGAAGGGCTGGCTCCCAAAGGAGCTTCAACTGCACTTGCCCTGGCTCTGGACCT
 CAAGGGCTTGAACGAGAACACTGAAAGGCAACAGAGACAAGAATGGACGGAAACGGGCCC
 TGCGGCTGCCAAAGGCTGTAGGGAGCAGGCCACGCAAGGGAAACAGCACCCCTTGGCAAT
 GATGTGCGCACGGCCTACAGCTCTGGCCCGCATTTACAGCATGAGAGGCCGAGCA
 GGGCTTGGACCTGGCAGCCACCCGAGAGGCTAATTTCTGAGGATGCGTCCATACAG
 GCAGCGCCCTCTGGCCCGAGCTACAGAGGATCGTGGGAACAGATCCAGCGAACGAG
 GCTGGTGAGCAGCTACTGAGGCATTTGGAGGCTACATTCAGCAACGTGGCACGAAA
 TGTGAAGAGGACCTATGAGGCTTCTGTCACGGCAACATGATTCTTGAG
 TTGACATCTGACAAAGCTGAACCTCACGGTGCCCAGGTGCAAGGTTGAGGACATT
 CAGGAAGAGCTCCAAGGGAGCTGGAGTCTCGGTGTCTCCCACTGACACCTTCAA
 GCCACCAAGAGAAAAAGAAGGCCCGTGGTGGAGGCTGACCAACCCGGAGGACTACCCAC
 TCACGGCACAAACAGAGGCCAGGGCTGAGAGGGAAACCTCATCCAGCACGGAGGAGA
 CACCCCGATGAGCCTGGACAGTTGCTGTTGCCCTGGTCAATTACCGGACCCCTGGG
 TCAGCTGCTGCCATGGACACTATGACCCCGACCATCGCAGCCTCCGACTGCCAACCGGC
 CTGTCATCAACACCCCCGGTGGTGGAGTGTCTGAGTGGTACAGTGAGGGAAACCCCACTCCCC
 AGCTCTGCAAGGGCTATCTGTTGGAGTCTCCCTGTTGGAGAGCAGGGAGGAAACGAAG
 CAAACCTGCTGTGTTCTGGAAACCAACTCCCTGACACTGGTGGAGTGGAGGGTGGT
 CAGCCAAGGGCTGTGAACTTCTGTCAGGGACCCGACCGTCACTGGCCAGTGCAGC
 CATTGGCCAGCTGCGCGGTGCTATGGACATTCCAGACGTGAGGACACGGGGAGGTTCT

FIGURE 9 CONTINUED

12/17

CCCCCTGAAGATCATCACCTATGCCGCCCTGTCCTTGTCTTGGTGGCCCTCTGGTGG
CTTTGTCCTCTCGCTCGTCCGACACTCGCCTCAACCTGCACAGCATCCCACAA
GAACCTATCCACGCTGTTCTCTCCAGCTCATCTTCATGGTCGGCATCAACCAGAC
TGAGAACCCGTTCTGCACAGTGGTCGCCATCCTCCTGCACATACGTCTCCATGGGCA
CCTTCGCCCTGGACCCCTTGAGAGAACTTGCACTGCTACCGCATGCTGACAGAAGTCGC
AACATCGACACTGGGCCATGGCGTCTACCAACTGGTGGCTGGGCATCCCTGCCAT
TGTACAGGACTGGCTGGACCTCAGGGCTATGGAAACCCCTGACTCTGCT
GGCTGTCCTCAGGATCTGGCTGGAGCTTGGCTGGCCTGTGGAAACGGTTATA
ATCATCAACACAGTCATCTTGTCTGCAAGGTTCTGCAAAAGAAAGACCA
TTATTATGAAAGAAAGGGGGTTGTCATGCTGAGGACGGCCCTCTCTGCTGCTGC
TCGTCAGTGCACCTGGCTGCTGGACTGCTGGCGTCAACAGTGACACTCTAGCTT
CACTACCTTGTGTCCTCAGCTGCTGCAGGGCATTTGTCCTCTGTTCTACTG
CGTGCACAAAGGGAGGTGCGGAAGCACCTGAGGGCGGTGCTGGCAGGGAAAGAGCTGC
AGCTGGATGACTCGGCCACCACTCGGCCACTGCTAACGCCCTCCCTCAACTGCAAC
AACACCTACAGCGAAGGGTCCAGACATGCTCCGACCCCTGGCAGTCCACAGCCTC
TCTGGACAGTACACCAAGGGATGAAGGGTCCAGAAACTCAGTGTGCTCTGGCCAG
CCCGTGGTAACCATGGAGAACAGATGCACTTCATCCCTAGGAACCTCAAAAAAGCT
CACGGCCCTGACTGACTCTGACAGTGAGCTGTCCTGGACGAGCACAGTAGTTCTA
CGCCTTCACACACATGGACAGCGAGGATGATGGCGAGAGGCTGAAGAACAAATGGA
ATCCGGCTGGGGCCCGCCATAGCACCCCAAAGCAGATGCTCTGGCAACCACGTC
CCAGCTGGCTGGCCGAGAGAGCTGGCTGGAGTGACAGTGAGGAGTTGGACACTGA
GCCCCACCTGAAGGTGAGAACAGGTCAGCGTGGAGTTACCCGGCAGGGCGCAGGGCAA
TCACTGTGGGACCCGGCCCTGACCCGGAAAGTGGGCTTGCCAAAGGCACTGGCG
TGCTTAGCAGCCAGCCTCAGAGCAGCGGAAAGGCACTCTGAAAACAAAGTCACCTAC
CCGCGGCCATTGCCAGAGCAGCCTGAAGTCCCGCTGCCAGAGAACGCTGGCTGATT
GTGAGCAGAGCCCCACATCCTCCCGACATCCTCCCTGGCTCTGGCGATGGTGTCCAT
GCCACTGACTGTGTCATTACCATCAAGACTCCGAGGAGGGAGCAGGCCGTGAGCATCT
CAATGGGTGGCATGAATGTACGACAGGGAGTGCCAGGCCAACGGTCTGACTCAGA
GAAACCATGAGGCACAGTCACCCACAGACTGCCGGTAAGGCCCTCAGACCTTGAAGC
CTGCGTGGACTGCTGCTATGGACAAGCAGGACCCCTTGACCCCTAGGGTCCCTGCCATA
GCAGCTGGCTTACGCAAGCGTCCAGACGGGAGGCCCTGACCCCTAGGAGCTCAG
GGCCAGACTCTGACAAGTGCACAGGCCACAGATGTCCTCAGAGGGAGACGTGACTT
CATTAAGGCTGGCATGCTCCGCTTAGACTGAAGACAGAATCCAACCATGTCGAC
AAGAGGCCATTGAGCCAGAGCTGGACTGGTGAATCATTGTCACGGCCCTCAACTGT
CCCGCAGGCCCTCCTTGTGGTACAAGCCCATCACCACAGCCTAGGGTGCCTTGCA
ACGGCACCCCTGGGTTAAATTGCTGTCATAAAATGAAATAGATATAATCTCTCCC
TGGACTTGGGAGAAGATGGGAGGCTGTGACTGCTTACACTGCTTGAETCTGCA
TGGAGAGGCCATGAATGGCATCTCACTCTATTGCCAAGGAAGCTTGACAGTTGACT
TGAATCTGGAAATGAGTCACTCAGCTGGTCCAGTGCCAGGTAGGGAGCATATGGGCT
GTGAAGTTGACACAGCTCTGCACTGCTTCCGGCCAGGGAGTGTTGGTGTCTGGCTGG
CCTTGGAGCACACGGCACATGTGTCGCTATGTCCTCAGCTCATGGATCCAGGT
CAGGGCTCACCTGTGAGGAGTGGGCCATAGCTATGATATGAAACTCTGACCCCTG
CCACCCCCCAGCCCCCCCCCCCCCCCCCTCTGATGCACTGAGGCGACTCTGGAGCCT
TTCCCGAGTCAGCCCAAGTGGCAGCGGGAGGACTCTGCTCTGCTCCATAAGCTCTCA
GAGTGCCTAGTGGCTGATTCTGAGGGAGGCCAGGGCTTCCAAAGCCTGTCACGGTCC
TTCTGCCAATGTTGACTTTTCAACTCCAGGGCTCTAGGCAGGCAGGTGCTCCA
CCCTGAAATCTGACCATGACATCACTTGCTTCAATGACCAATTGCAAGAAC
AAAGCCAGAGTGGCCCTTCAATGGTACCCATTCTTGGAAATCTCACACCAAGG
ACCTGTGACCGCCACTGAGAGGCCAGGGTGCAGGCCAGGCAAGGGATGGGAGCCTGGA
GTTACTCAGCACGTTTACTCAGCATGGACTGTTGACTGAATCAGGTCCCCAAAGTAC
ATGGGTGACAGTCGCTCGGAATGGAGAACCTCAGGGAGGGCTCAAAGGGCAGGAC
TTCTCAGGCTGTCAGAAGTGCACAGGACTGCTAGTGACTGACTGGTGGAGAT
GAAAGCTTACAAGACATGGCTGGGGTACGCACTGCAAGAGGCTGACGGAGGCCA
GGTAGCCCAAGGATGGCAAGGATACAGAGTGACCTAGCAAGGGGAGCTTCACTG
GTGGTACAGCACCGTACAACCTCCGAACCCACGCCACCTCAGAAGGTGAAGTTTGA
TTCGATCACAATTAGCAAACAAACCCCTGCACTTAAACTGTTTCTGACCTA
AGACTTTCTGACCGTAAGACATGGAGATTTAACAGGTGTATTACTGTCAGC

FIGURE 9 CONTINUED

13/17

ACTGGATGGCAACACAGGTGAGATGATGCATCTATAATAAAATTAAAGATTTGGATTTG
TAAA

FIGURE 9 CONTINUED

14 / 17

FIG. 10

	5	10	15	20	25	30	35	40	45	50
	NSGLM	EKLKQ	IEEQT	KKAAQ	ELEEQ	TRRAL	ELEQE	RKRAG	TAVIE	LRAHD
	55	60	65	70	75	80	85	90	95	100
C01	PDEGD	AGRLS	YQMEA	LFDER	SNGYF	LIDAA	TGAVT	TARSL	DREIK	DTHVL
	105	110	115	120	125	130	135	140	145	150
	KVSAV	DHGSP	RRSAA	TYLTV	TVSDT	NDHSP	VFEQS	EYRER	IRENL	EVGYE
	155	160	165	170	175	180	185	190	195	200
C02	VLTIR	ATDGD	APSNA	NMRYR	LLEGA	GGVFE	IDARS	GVVRT	RAVVD	REEAA
	205	210	215	220	225	230	235	240	245	250
	EYQLL	VEAND	QGRNP	GPLSA	SATVH	IVVED	ENDNY	PQFSE	KRYVV	QVPED
	255	260	265	270	275	280	285	290	295	300
C03	VAVNT	AVLRV	QATDR	DQGQN	AAIHY	SIVSG	NLKQQ	FYLHS	LSGSL	DVINP
	305	310	315	320	325	330	335	340	345	350
	LDFEA	IREYT	LRIKA	QDGGR	PPLIN	SSGLV	SVQVL	DVNDN	APIFV	SSPFQ
	355	360	365	370	375	380	385	390	395	400
	AAVLE	NVPLG	HSVLH	IQAVD	ADAGE	NARLQ	YRLVD	TASTI	VGGSS	VDSEN
C04	405	410	415	420	425	430	435	440	445	450
	PASAP	DFPFQ	IHNSS	GWITV	CAELD	REEVE	HYSFG	VEAVD	HGSPA	MSSSA
	455	460	465	470	475	480	485	490	495	500
	SVSIT	VLDVN	DNDPM	FTQPV	YELRL	NEDAA	VGSSV	LTLRA	RDRDA	NSVIT
	505	510	515	520	525	530	535	540	545	550
C05	YQLTG	GNTRN	RFALS	SQSGG	GLITL	ALPLD	YKQER	QYVLA	VTASD	GTRSH
	555	560	565	570	575	580	585	590	595	600
	TAQVF	INVTD	ANTHR	PVQFS	SHYTV	SVSED	RPVGT	SIATI	SATDE	DTGEN
	605	610	615	620	625	630	635	640	645	650
C06	ARITY	VLEDV	VPQFR	IDPDT	GTYYT	MTELD	YEDQA	AYTLA	ITAQD	NGIPQ
	655	660	665	670	675	680	685	690	695	700
	KSDTT	SLEIL	ILDAN	DNAPR	FLRDF	YQGSV	FEDAP	PSTSV	LQVSA	TDRDS
	705	710	715	720	725	730	735	740	745	750
C07	GPNGR	LLYTF	QCGDD	GDGDF	YIEPT	SGVIR	TQRRL	DRENV	AVYNL	WALAV
	755	760	765	770	775	780	785	790	795	800
	DRGSP	NPLSA	SVGIQ	VSVLD	INDNP	PVFEK	DELEL	FVEEN	SPVGS	VVARI

FIG. 10 CONTINUED

15 / 17

	805	810	815	820	825	830	835	840	845	850
008	RANDP	DEGPN	AQIIY	QIVEG	NVPEV	FQDL	LSGDL	RALVE	LDFEV	RRDYM
	855	860	865	870	875	880	885	890	895	900
	LUVQA	TSAPL	VS RAT	VHIRL	LDQND	NPP EL	PDFQI	LFN NY	VTKNS	NSFPS
	905	910	915	920	925	930	935	940	945	950
009	GVIGR	IPAHD	PDLSD	SLNYT	FLQGN	ELSLL	LLDPA	TGELO	LSRDL	DNNRP
	955	960	965	970	975	980	985	990	995	1000
	LEALM	EVSVS	DGIHS	VATALC	TLRVT	IITDD	MLTNS	ITVRL	ENMSQ	EKFLS
	1005	1010	1015	1020	1025	1030	1035	1040	1045	1050
	PLSL	FVEGV	ATVLS	TTKDD	IFVFN	IQNDT	DVSSN	ILNVT	FSALL	PGGTR
	1055	1060	1065	1070	1075	1080	1085	1090	1095	1100
	GRFFF	SEDLQ	EQIYL	NRTLL	TTISA	QRVLP	FDDNI	CLREP	CENYM	KCVSV
	1105	1110	1115	1120	1125	1130	1135	1140	1145	1150
	LRFDG	SAFFI	SSTTV	LFRPI	HPITG	LCRC	PPGFT	GDYCE	TEIDL	EGF1
	1155	1160	1165	1170	1175	1180	1185	1190	1195	1200
	CGANG	GCRSR	EGGYT	CECF	DFTGE	HCQVN	VRSGR	CASGV	CKNGG	TCVNL
	1205	1210	1215	1220	1225	1230	1235	1240	1245	1250
	LIGGF	HCVCP	FGEYE	HPYCE	VSTRS	FPPQS	FVTFR	GLRQR	FHFTV	SLAFA
	1255	1260	1265	1270	1275	1280	1285	1290	1295	1300
	TQDRN	ALLLY	NGRFN	EKHDF	IALEI	VEEQL	QLTFS	AGETT	TTVTP	QVPGG
	1305	1310	1315	1320	1325	1330	1335	1340	1345	1350
	VSDGR	WHSVL	VQYYN	KPNIG	HLGLP	HGPSC	EKVAV	VTVDD	CDAAV	AVHFG
	1355	1360	1365	1370	1375	1380	1385	1390	1395	1400
	SYVGN	YSCAA	QGTQS	GSKKS	LDLTG	PLL LG	GVPNL	PEDFP	VHSRQ	FVGCM
	1405	1410	1415	1420	1425	1430	1435	1440	1445	1450
	RNL SI	DGRIV	DMAAF	IANG	TRAGC	ASQRN	FCDGT	SCQNG	GTCVN	RWNTY
	1455	1460	1465	1470	1475	1480	1485	1490	1495	1500
	CECP	LRFGG	KNCEQ	AMPHP	QRFTG	ESVVL	WSDLD	ITISV	PWYLG	LMFRT
	1505	1510	1515	1520	1525	1530	1535	1540	1545	1550
	RKEDG	VLMEA	TAGTS	SRLHL	QILNS	YIRFE	VSYGP	SDVAS	MQLSK	SRITD
	1555	1560	1565	1570	1575	1580	1585	1590	1595	1600
	GGWHH	LLIEL	RSAKE	GKD IK	YLAVM	TLDYG	MDQST	VQIGN	QLPGL	KMRTI
	1605	1610	1615	1620	1625	1630	1635	1640	1645	1650
	VIGGV	TE DKV	SVRHG	FRGCM	QGVRM	GESST	NIATL	NMND	LKV RV	KDCCD

16/17

FIG. 10 CONTINUED

1655 1660 1665 1670 1675 1680 1685 1690 1695 1700
 EGF4 VEDPC ASSPC PPHRP CRDTW DSYSC EGF4 ICDRG YLEKK CVDAC LLNPC KHVG
 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750
 EGF5 LCALP NTPRG YSCEC GPGHY GQYCE NKVDL PCPKG WNGNR CVAPV TVLSA
 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800
 DIVERGED EGF
 KALIP TATRP MASAR RITTS PQPRI VAFPV TVSPR SHSRA CDMDT GQAC
 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850
 KPGVI GRQCN RCDNP FAEVT SLGCE VTIYNG CPRAF EAGIW WPQMK FGQPA
 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900
 AVLCP KGSVG NAVRH CSGEK GWLPP ELFNC TGSF VDLKA LNEKL NRNET
 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950
 RMDGN RSLPL AKALR NATQG NSTLF QNDVR TAYQL LARIL QHESR QQGFD
 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000
 LAATR EANFH EDVWH TGSAL LAPAT EASWE QIQRS EAGAA QLLRH FEAYF
 2005 2010 2015 2020 2025 2030 2035 2040 2045 2050
 SNVAR NVERT YLRPF VIVTA NMILA VDIFD KLNFT GAQVP RFEDI QEELP
 2055 2060 2065 2070 2075 2080 2085 2090 2095 2100
 RELES SVSFP ADTFK PPEKK EGPVV RLTNR RTTPV TAQPE PRAER ETSSS
 2105 2110 2115 2120 2125 2130 2135 2140 2145 2150
 RRPRH PDEPG QFAVA LUVYI RTLQG LLPEH YDPDH RSLRL PNRPV INTPV
 2155 2160 2165 2170 2175 2180 2185 2190 2195 2200
 VSAMV YSEGT PLPSS LQRPI LVEFS LLETE ERSKP VCVFW NHSLD TGGTG
 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250
 GWSAK GCELL SRNRT HVTCQ CSHSA SCAVL MDISR REHGE VLPLK IITYA
 TM1 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300
 ALSLS LVALL VAFVL LSLVR TLRSN LHSIP QEPHK ALFFS QLIFM VGINQ TM2
 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350
 TENPF LCTVV AILLH YVSMG TFAWT LVENL HVYRM LTEVR NIDTG PMAFY TM3
 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400
 TM4 HVYGN GIFAJ VTGLA VGLDP QGYGN PDFCW LSLQD TLIWS FAGPV GTVII TM5
 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450
 INTVI FVLSA KVSCQ RGHYI YERKG VVSMR RTAFL LLLLV TATWL LGLLA TM6
 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500
 TM7 NSDT LSFHY LFAAF SCLQG FVLL FYCVA NREVR KHLRA VLAGK KLQLD

17 / 17

2505 2510 2515 2520 2525 2530 2535 2540 2545 2550
DSATT RATLL TRSLN CNNTY SEGSR HAPHR PGQST ASLDS TTRDE GVQKL
2555 2560 2565 2570 2575 2580 2585 2590 2595 2600
SVSSG PAPGN HGEPD ASFIP RNSKK AHGPD SDSDS ELSLD EHSSS YASSH
2605 2610 2615 2620 2625 2630 2635 2640 2645 2650
TSDSE DDGGE AEDKW NPAGG PAHST PKADA LANHV PAGWP DESLA GSDSE
2655 2660 2665 2670 2675 2680 2685 2690 2695 2700
ELDTE PHLKV RPRSA WSYTG RRRAI TVATG PLTRK VGSWP SQWPC LAASP
2705
RSSCK AS

FIG. 10 CONTINUED

THIS PAGE BLANK (USPTO)